TAX COMPLIANCE: THE EFFECT OF THE TIMING OF PAYMENTS

By

RICHARD BENNETT DUSENBURY

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1989

ACKNOWLEDGMENTS

This thesis uses theories, methods, techniques, formats and procedures all of which are due to others. In addition, there are some whose personal involvement have facilitated and enhanced this study. These contributors are the faculty who generously took time to react to the project at various stages, including Jon Hamilton, Jack Kramer, and Gary McGill as well as Ph.D. students in accounting and marketing. Sandy Kramer must be acknowledged singly. She has been particularly helpful in the process from the development of the design to the review of the written report.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS ii

ABSTRACT v

CHAPTER

1 INTRODUCTION 1

Introduction 1
Background To Prepayment Arrangements 2
Timing Of Payments And The Time Value of Money . . 7
Taxpayer Prepayment Patterns, 1955-1985 9
The Research Question 15
Implications Of The Research 18

2 THEORETICAL BACKGROUND AND HYPOTHESES 20

Introduction 20
Theoretical Background:
The Sociological Tradition 20
Theoretical Background:
The Economics Tradition 28
Some New Theories Of Rational Choice 31
Hypotheses 37

3 REVIEW OF RELEVANT RESEARCH 39

Introduction 39
Risk Preference Reversal: Utility
Maximization Or Bias 39
Applications Of Prospect Theory. 44
Limitations Of The Existing Tax Research 50
New Directions For Research On The PR Effect . . . 53

4 THE RESEARCH DESIGN 56

Introduction 56
Variables And The Model 56

Page

The Hypotheses 61
Experimental Process And Procedures 64
Design Innovations 71
Manipulation Checks And Controls 79
Implications Of Results For Utility
Representations 84

5 RESEARCH RESULTS 93

Introduction 93
The Hypotheses 94
Manipulation Checks And Controls 102
Implications Of Results For Utility
Representations 115

6 SUMMARY AND CONCLUSION 127

Summary 127
Limitations 132
Conclusion 133

APPENDIX: EXPERIMENTAL INSTRUMENT 141

REFERENCES 154

BIOGRAPHICAL SKETCH 162

Abstract Of Dissertation Presented To The Graduate School Of The University of Florida In Partial Fulfillment Of The Requirements For The Degree Of Doctor Of Philosophy

TAX COMPLIANCE: THE EFFECT OF THE TIMING OF PAYMENTS

By

Richard Dusenbury

December, 1989

Chairperson: Sandra Kramer
Major Department: Accounting

Income tax law requires substantial prepayment of the tax liability reported in the tax return. A payment accompanies the tax return only if the prepayments are less than the liability reported. The research question is: Does the timing of payments, which produces a payment due or an overpayment, affect tax compliance? An important finding is that the timing of payments affects the amount of tax liability reported. Taxpayer subjects report significantly lower liabilities in payment-due than in refund tax filings.

Cash incentives endow the experimental task with risks, costs, and benefits, which are the attributes of decision making under uncertainty. The empirical results are used to evaluate the risk preferences of taxpayers. The finding is that all tax filing choices are risk-averse but that riskier

reporting positions are chosen in payment-due than in refund filings.

Two nontax cases, a purchase of health insurance and selection of a gamble, have parameters identical to two of the tax cases. The nontax cases are used to measure the willingness of taxpayers to take risky tax filing positions relative to their willingness to take risk in other decisions. The finding is that the willingness to accept risk is slightly lower in the health insurance case than in the payment-due tax case and is much higher in the gamble than in all of the tax cases.

The implications of the results for utility function and value function representations of preference orderings are assessed. The empirical results in the tax cases are not incompatible with a risk-averse, concave utility function or an s-shaped value function. The significant context differences are incompatible with a single functional representation.

CHAPTER 1

INTRODUCTION

Introduction

The federal income tax is imposed on individuals in installments, and substantial prepayment of the annual liability is required. The individual tax filings are the formal accounting each taxpayer makes to measure, and report, his actual tax liability for the year. This accounting results in a settlement between the taxpayer and the government. If the taxpayer has prepaid more than the liability he reports, he may either have the excess amount refunded or have the excess amount applied to his tax liability for the following year. If the taxpayer has prepaid less than the liability he reports, then an additional payment is required. The additional payment is submitted along with the tax filing.

The present arrangements result in payments of federal income taxes by individuals which are not contemporaneous with the accounting the individual uses to determine his annual liability. The tax filing, in which the tax liability is measured, is linked with the settlement amount. In the most general sense, the problem examined in this study

involves some of the effects of administering federal income taxes using payment arrangements in which the settlement amount, which is either a refund or a payment due, is impounded in the tax filing.

Background To Prepayment Arrangements

An important mechanism for the prepayment of individual income taxes is the withholding system. In general, employers are required to withhold portions of salaries paid to employees. Withholding is mandated in Section 3402 of the Internal Revenue Code. This provision directs the Treasury Department to adopt rules and regulations for implementing the statute. The Treasury Department has adopted Treasury Regulations containing detailed charts and schedules for withholding of taxes from salaries and wages. Withholding is required on some interest and dividend income under some circumstances, and the extension of withholding is continually debated. For example, the extension of withholding to all dividends and interest was part of President Kennedy's tax proposal in 1962. The proposal was intended to be a revenue increasing provision to partly offset the investment tax credit which was also proposed (Witte, 1985, p. 156).

We know surprisingly little about the effects of using this particular system for the payment of income tax liabilities. Prepayment in installments is not typical of other types of taxes such as personal and real property

taxes. Typically property taxes are due at the end of a period, and discounts are available for payment earlier than the due date. The withholding system for federal income taxes has been in place for almost 50 years, yet it has been relatively unexamined.

Withholding was implemented during World War II to insure collection of income taxes when the taxes were imposed on individuals with modest incomes. Between the implementation of the modern federal income tax in 1913 and 1941, the tax rates had varied from modest to confiscatory levels. Regardless of the rates, the tax had been imposed on very few individuals (Doernberg and McChesney, 1987, pp. 914-915). The war effort generated terrific revenue demands, and both the rates and the tax base were greatly expanded. Under new tax provisions enacted in 1942, 40,000,000 tax returns were due (50,000,000 if the 5% Victory tax on all income over $624 is counted). This was at a time when the U.S. population was about 125,000,000. Tax liabilities at 24% of taxable income were imposed on single taxpayers with incomes greater than $600. In the first year in which the new tax was imposed, collections were dismal. In addition, the wartime demand was stimulating economic growth, and inflation was a by-product of the economic expansion. Withholding was already in place for the recently implemented social security tax system. Withholding for federal income taxes was implemented to remove purchasing power from the employed and to collect

substantial taxes imposed on individuals with modest incomes (Witte, 1985, pp. 114-125).

Withholding was credited for the success of the collection of the substantial taxes levied on persons with modest incomes. It was assumed that small payments deducted periodically were valued differently from one-time, out-of-pocket payments by all taxpayers, and the forced savings was thought to be particularly critical for modest income taxpayers (Lindholm, 1948, pp. 269-270; and Homuth, 1948, p. 346). Without withholding, it was presumed that the low income taxpayer would spend his entire income and be unable to pay his tax liability. Thus the wartime experience with withholding yielded a common wisdom that payment in small installments was essential for imposing income taxes on persons with low incomes.

Another specific feature of the withholding arrangements used in the United States is worth noting. Using IRS rates and schedules, employers compute withholding as though the salary on which withholding is made for that pay period is the same amount that will be earned for all pay periods in a year. This can often cause withholding in excess of the annual tax liability. For example, if one has wages for only six months, the withholding over those six months will be at a rate which assumes the taxable income from wages will be twice as high as it turns out to be. An alternative scheme might be one in which the withheld amounts are based on

accumulated year-to-date earnings rather than on the earnings within each pay period (Kay and King, 1986, pp. 28-32). The bias toward overwithholding is recognized by the Internal Revenue Service (IRS). The policy has been to use withholding rates and schedules which will yield prepayments which are larger than the taxpayer's liability (Sheppard, 1987, p. 439). The IRS has been blamed for the extent of taxpayer overpayments, and the IRS was directed in the Tax Reform Act of 1986 (TRA86) to change the overwithholding bias (Jones, 1986, p. 971).

The IRS did attempt to eliminate the bias toward overwithholding. A withholding form, called Form W-4, is completed by each employee and is filed with his employer. Employers use the information about dependents and filing status on the Form W-4 in determining the amount of withholding. Pursuant to the passage of TRA86, the IRS created a worksheet for completing the new Form W-4. The new Form W-4 was supposed to produce prepayments through withholding which varied little from the actual liability (Uhlfeder, 1986, p. 703). This worksheet and new Form W-4 was issued in November of 1986. Since all wage earning taxpayers were involved, many individuals were affected. The response was certain and swift. Politicians received complaints from constituents that the worksheet was too complicated. Within two months, both House and Senate committees held hearings on the new forms. The Commissioner

of the IRS continued to justify the complexity of the new procedure on the grounds that the complexity was necessary to increase the accuracy of taxpayer prepayments (Teuber, 1987, pp. 633-635). The upshot of this brouhaha was the creation of a new worksheet and Form W-4 which was issued in March of 1987. The new procedures, although simpler, partially restored the bias toward overpayment which typified withholding before 1986 (Johnson, 1987, pp. 946-947).

The withholding system constitutes requirements imposed on employers, not on taxpayers. In general, an employer may become financially responsible for income taxes owed by an employee if the liability is from salaries or wages on which the employer neglected to withhold. The requirement on the taxpayer to prepay taxes comes from a different statute, Section 6654 of the Internal Revenue Code. This provision requires that tax payments be prepaid in installments. This provision requires quarterly installments such that, in general, at least 90% of the liability declared in the tax filing be prepaid well before the filing deadline. If the taxpayer does not meet these requirements, an interest penalty on underpayment of installments is imposed. To avoid penalties when withholding produces insufficient prepayments, taxpayers must make quarterly installment payments, called estimated payments. For nearly all taxpayers, the prepayments through withholding have produced enough prepayment so that the taxpayer is in compliance with the requirement that taxes

be prepaid in installments. In 1985 two percent of all tax returns had a predetermined estimated tax penalty (IRS, 1985, p. 36). In 1977, the most recent year for which data is available, 86% of all individual returns for which there was a tax liability had prepayments only through withholding (IRS, 1977, pp. 129-134).

Many aspects of the tax system have been continually reformed, but the arrangements for the payment of income tax liabilities has been very stable since the 1940s. Currently, lower income individuals do participate in the income tax system. About 100,000,000 individual tax returns are filed each year from a population of about 250,000,000. While about 20,000,000 of these filings are made only to claim refunds of withheld salaries when no liability exists, many of the remaining 80,000,000 returns are joint filings from married couples. For most taxpayers, the prepayment of income taxes still occurs through withholding. The current arrangements for the payment of income tax liabilities have been relatively unexamined in empirical tax research.

Timing Of Payments And The Time Value Of Money

Until recently, there has been no theoretical basis for evaluating the effects of the timing of payments except for consideration of the time value of money. Instead, the entire focus has been on the amount of the liability declared without consideration of the timing of payments. One traditional source for theories of taxpaying behavior is the

work on criminal behavior in general and white collar crime in particular (Jackson and Milliron, 1986; Sutherland and Cressey, 1970; and Lemert, 1964). The variables of interest from this perspective are attitudes and experiences associated with law-abiding and law-breaking behavior. The amount of the liability reported, from this perspective, is unaffected by the timing of the tax payments. Another theory source is the work done by economists in the field of public finance. A product of this work is what has become known as the agency model (Mirrlees, 1971 and 1976; Holmstrom, 1979; Allingham and Sandmo, 1972; Becker, 1968; and Reinganum and Wilde, 1985 and 1988). This approach stresses the risk and penalties for understating the liability. Like the sociological approach, this approach omits any consideration of the timing of payments. Only the time value of money offers any direct link between mainstream economic theory and the timing of the payments.

The notion of discounting cash flows according to the timing of the cash flows was applied in both macroeconomic and microeconomic theory and analysis by 1950 (Duesenberry, 1949; Fisher, 1930; Friedman, 1957; Bierman and Smidt, 1971; Dean, 1951; and Lutz, 1951). The time value of money is commonly considered in tax choices, but the application is through the timing of the liability, not the payment of the liability. For example, one may decide not to sell an appreciated capital asset so as to delay recognition of

income. This may be a strategy to delay the cash outflow in paying the tax liability, but the payment is delayed by delaying the liability. The continual debate regarding effects of capital gain taxation on revenues (do higher rates reduce revenue?) are an indication of widespread awareness that the timing of a tax liability is important in taxpayers' choices (Gravelle and Lindsey, 1988; Toder and Ozanne, 1988; and Kotlarski, 1988).

On the other hand, there is no body of literature on the application of the time value of money to income tax payments for a given liability, probably because the application is so straightforward. Specifically, the discounting of cash flows means that an optimal choice consists of delaying cash outflows and speeding up cash inflows. Thus, the optimal payment strategy, given the time value of money, is to prepay as little as possible and to make the payment due with the tax filing as large as possible. This strategy appears to be rejected by the overwhelming majority of taxpayers.

Taxpayer Prepayment Patterns, 1955-1985

Most individual taxpayers seem to hold a strong preference for being overpaid at the time of filing even though overpayment would seem to be uneconomical. In 1981, a typical year, there were 70,000,000 individual tax returns with overpayments. This amounted to 75% of all individual returns. The average overpayment in 1981 was $809. Comparing the overpayment to the average tax liability for all

individual taxpayers provides a basis for visualizing the size of the overpayments. In 1981 the average tax liability was $3,034, so the average overpayment was 27% of the average liability. This means that, on average, U.S. taxpayers prepaid amounts equal to about 125% of their tax liabilities (IRS, 1985).

One reason for overpayments is the bias toward overpayment built into the withholding system. This means that individuals who are not employed for the entire year have withholding taken out at a rate which assumes they are employed all year. In many cases, such taxpayers have no liability at all, and they must file returns solely to have withheld amounts refunded. The bias in the withholding procedures does affect low-income taxpayers more heavily than other taxpayers, but the extent of overpayments at all income levels is still high. Table 1-1 shows data for 1981 by several Adjusted Gross Income (AGI) levels. Only at AGI levels greater than $20,000 did the overpayment fall below 39% of the average liability. For taxpayers with AGI amounts greater than $20,000, the overpayments still exceed 21% of the average liability, and the returns with overpayments constitute 64% of all returns. The variation by AGI level in 1981 is typical of other years.

The pattern of substantial overpayments by most taxpayers has been extraordinarily stable. Table 1-2 shows the same data as were described for 1981 in total, but it

Table 1-1
Overpayment Data For 1981 By
Adjusted Gross Income Levels

Adjusted Gross Income	Number Of Returns With Over-Payments	Overpay-ment Returns As A Per-cent Of All Tax Returns	Aver-age Over-pay-ment	Average Tax For All Tax-payers	Average Overpay-ment As Percent Of Aver-age Tax, All Tax-payers
<$5,000	15,410,867	91%	$304	$33	933%
$5-10,000	13,864,418	81%	$501	$429	117%
$10-15,00	10,887,070	77%	$581	$1,199	48%
$15-20,00	8,132,715	74%	$805	$2,053	39%
>$20,000	21,733,671	64%	1,479	$6,958	21%
Total	70,028,741	75%	$809	$3,034	27%

includes various years from 1955 to 1976 and each year from 1976 to 1985. During this period of 30 years, there have been marked changes in tax rates and major tax law changes. There also have been variations in economic conditions from business cycles, and wide variations in interest rates and inflation, yet the data on overpayments reflect a very stable pattern. Since 1970 between 75% and 80% of all returns have overpayments. The average overpayment has steadily increased, and it has always been between 24% and 31% of the average liability in each year.

The time value of money does not seem to be a factor in this pattern. If taxpayers are responsive to the time value of money, there should be an inverse relationship between interest rates and overpayment amounts. Consider the time series of overpayments and interest rates in Table 1-3. It is apparent from observation that the refund amounts steadily increase while the interest rates fluctuate. Over this period of 30 years, the interest rate trend has been upward, so that overpayments and nominal interest rates are positively, and highly, correlated (Pearson correlation coefficient of .84). However, a regression using first differences from this span of 30 years produces a model with an R-square of .01 and a statistically insignificant relationship between the changes in interest rates and the changes in the refunds. Between 1976 to 1985 interest-bearing checking accounts and high interest rates enhanced the value

Table 1-2
Overpayment Data For 1955 - 1985

Year	Number Of Returns With Over-Payments	Overpay-ment Returns As A Per-cent Of All Tax Returns	Aver-age Over-pay-ment	Average Tax For All Tax-payers	Average Overpay-ment As Percent Of Aver-age Tax, All Tax-payers
1955	34,045,137	65%	$92	$562	16%
1957	35,923,474	66%	$105	$631	17%
1960	39,372,379	68%	$145	$686	21%
1963	41,403,733	68%	$168	$795	21%
1966	49,380,446	74%	$174	$835	21%
1969	53,075,526	75%	$246	$1,220	20%
1970	55,273,385	77%	$285	$1,168	24%
1974	65,764,063	81%	$412	$1,548	27%
1976	64,976,027	79%	$482	$1,937	25%
1977	66,031,939	79%	$511	$2,111	24%
1978	63,775,981	75%	$518	$2,181	24%
1979	71,433,691	79%	$641	$2,626	24%
1980	69,868,451	76%	$708	$2,711	26%
1981	70,028,741	75%	$809	$3,034	27%
1982	72,425,701	78%	$876	$2,976	29%
1983	74,993,969	80%	$906	$2,907	31%
1984	75,563,505	78%	$929	$3,569	26%
1985	77,424,003	78%	$984	$3,798	26%

Table 1-3
Overpayments and Interest Rates, 1955-1985

YEAR	AVERAGE OVERPAYMENT (In Dollars)	NOMINAL 3 MOS. T. BILL RATE (In Percent)
1955	92	1.75
1956	110	2.66
1957	105	3.26
1958	128	1.84
1959	133	3.41
1960	145	2.95
1961	150	2.38
1962	161	2.78
1963	168	3.18
1964	150	3.55
1965	154	3.95
1966	174	4.88
1967	199	4.33
1968	218	5.35
1969	246	6.69
1970	285	6.44
1971	279	4.34
1972	371	4.07
1973	388	7.03
1974	412	7.87
1975	460	5.82
1976	482	5.00
1977	511	5.27
1978	518	7.22
1979	641	10.04
1980	708	11.61
1981	809	14.08
1982	876	10.72
1983	906	8.62
1984	929	9.57
1985	984	7.49

benefits of slower tax payments for average taxpayers, but the relationship between the change in interest rates and the change in overpayments is statistically insignificant for this decade as well.

It appears that taxpayers have responded to an income tax system with a broad base, payments due in installments, and prepayments through withholding with a strong preference for overpayment of taxes. Otherwise some other influence contrary to taxpayer preferences must produce this kind of stable pattern. An argument that taxpayers keep trying to prepay at exactly 100% and fail in exactly the same way year after year seems ludicrous. The public outcry over the new Form W-4 and worksheet in 1986-1987 also suggests that taxpayers were neither dissatisfied with their existing prepayments, nor were they enamored of changes designed to reduce their overpayments. It appears that most taxpayers prefer overpayment instead of underpayment. The typical tax filing is a filing for a refund.

The Research Question

The research question involves the relationship between the timing of payments and the amount of tax liability reported. In mainstream theory, the relationship between timing of payments and the declared liability is pictured as one in which the liability might affect the timing of payments. For example, taxpayers may delay payments by delaying liabilities. The new concern is that a causal

relationship between the liability declared and the timing of payments might flow in the other direction; the timing of tax payments may affect the amount of the liability. Anecdotes of those involved in tax administration, tax practice, and tax filing illustrate the importance of the timing of tax payments. For example, tax practitioners claim that clients have markedly different attitudes about a filing with a payment due as compared to a filing when a refund is obtained. If taxpayers behave differently in overpayment and payment-due filings, then practitioners should treat clients with overpayments and clients with payments due differently. In fact, there is some empirical evidence that practitioners make different recommendations to clients who are significantly overpaid at filing than they make to clients who are significantly underpaid at filing (Duncan, LaRue and Reckers, 1989).

The state of the empirical research on the question has not always been well focussed around the issue of the timing of the payments. To date, the exploration of the effect of the timing of income tax payments on the amount of tax liability reported has involved a quest for a so-called payment effect. A payment effect, hereafter termed a PR effect, is the effect of the settlement amount (payment due or refund) on the taxpayer's choice of the taxable income reported on the tax return. This formulation compresses all timing variation into a single difference at settlement.

That is, the taxpayer has either prepaid more than the liability or less than the liability at the time that he prepares the accounting through which he declares his liability. The PR effect defines the effect of timing differences in terms of a significant objective function, namely the amount of taxable income reported by the taxpayer. The issue is not merely whether taxpayers have a preference between being overpaid or underpaid at the time of filing, but whether any such preference has an effect on the amount of tax liability the taxpayer reports. Any event that affects the amount of the liability also alters the payment due or overpayment in a tax filing. For example, a change in the tax law may change a tax liability such that an additional payment is owed even though, absent the change, the taxpayer was overpaid at the time of filing. The current research often confounds these variances between the expected and actual liability with any PR effect. The current research is described in Chapter 3.

The research question addressed here is: Is there a PR effect on the amounts taxpayers report in their tax filings? For our purposes, the PR effect will exclude variations in payments and refunds produced by events affecting the amount of the tax liability. Therefore, if two tax filings are identical but for the fact that one has an additional payment due and the other is a filing for a refund, then the two filings differ only in prepayments. Chapter 4 describes an

experiment conducted to collect evidence on this question. Sixty-five experienced taxpayers completed three tax filing cases plus two nontax spending choices. Subjects received cash incentives proportional to the ending balance of a fund used to make payments in the cases. Chapter 5 reports the research results.

Implications Of The Research

If a PR effect exists, there are potentially wide-ranging implications for tax administration and tax practice. Virtually every tax filing would be affected. Consequently, even if a PR effect were relatively insignificant in comparison to other determinants of compliance, the aggregate impact on the tax system could be huge. If a PR effect exists, taxpayer satisfaction, taxpayer compliance, and the revenue attained under a tax regime will vary with the withholding requirements and prepayment practices of taxpayers. Therefore, a PR effect has implications for tax policy regarding withholding, taxpayer selection of withholding and estimated payment strategies, and even a tax-return preparer's relationship with his clients. One example of this latter situation occurs if a tax client is likely to be more satisfied with unaggressive choices (ones safer and often less expensive for the tax practitioner) when there is a refund to be obtained than when there is an additional payment due. Though the research question has potential policy implications, the objective is to examine the

empirical question of whether or not an effect exists. Chapter 6 includes the limitations, a summary of the findings, and the conclusions from the experiment conducted to collect evidence on a PR effect.

CHAPTER 2

THEORETICAL BACKGROUND AND HYPOTHESES

Introduction

The two sources for most models of tax compliance are rational choice models from economics and theories of legally abiding behavior from other social sciences. Neither source provides a basis for believing that the timing of tax payments is a factor in tax filing decisions. Consequently, even though researchers have evaluated a substantial variety and number of variables which may affect tax compliance, the timing of tax payments is not one of those variables. Newer approaches to rational choice, however, provide a basis for predicting a PR effect.

Theoretical Background: The Sociological Tradition

Sociological theory is one source for models of tax compliance. During the 1920s citizens in many cities and states felt that public order had been completely destroyed by criminal behavior generally and gangster activity specifically. The fear for public order was reflected in intense public involvement in crime control. For example, during the 1920s there were dozens of city and state crime commissions, and crime surveys were conducted in every

sizeable city in the country. Finally a national crime commission called the Wickersham Commission was established in 1929 (National Commission on Law Observance and Enforcement, 1931; Kuhlman, 1929; and Conner, 1930).

Before 1930, a scientific explanation of vandalism, youth gangs, juvenile delinquency, and organized crime was in high demand. Sociologists lacked a theory of illegal conduct, but in this context of high demand, they could only sustain a claim to expertise by providing empirical research. Sociologists consequently proceeded on to the empirical research, and, in the absence of theory, they emphasized research methods. They claimed to have two methods, the case study and the survey, to collect information about deviant behavior. The titles are enough to indicate whether the case or survey method was in use. Exemplary studies were Six Boys in Trouble, a Sociological Casebook (Reckless, 1929), Delinquency Areas (Shaw, Zorbaugh, McKay, and Cottrell, 1929), The Unadjusted Girl (Thomas, 1923), The Gang (Thrasher, 1927), The Gold Coast and the Slum (Zorbaugh, 1929), and The Jack-Roller (Shaw, 1930). This emphasis on research methodology was undertaken with the hope that the information would ultimately disclose general patterns that could be captured inductively in theory (Odum and Jocher, 1929, p. 246).

The result of the extensive descriptive research was an explanation of disobedient behavior which probably does

not deserve to be called a theory even though it was designated multiple factor theory. The notion was that deviant behavior was the result of many factors such as lack of playgrounds, poverty, immigration, slums, broken homes, the search for excitement, material gain, and so on. All the factors had to be taken into account to explain illegal conduct.

The multiple factor approach had an empirical problem and a theoretical problem. The theoretical problem was that the relationships among the factors were unspecified. The empirical problem was that each factor, consequently, was a contributing cause to disobedient conduct. Since no theory was available to explain how the factors interacted, or how they conditioned one another, each factor was logically considered to have some independent impact. This conclusion, however, was patently absurd, because when people were subjected to identical causal factors, some behaved obediently and some did not. Obedient children and criminals came from the same schools, ethnic groups, neighborhoods, and even families. The multiple factor theory negated this absurdity. According to this theory, no one factor could be considered in isolation from others. Instead, the multiple motives and characteristics of disobedient individuals, whether juvenile delinquents or organized mobsters, all had to be taken into account to explain their noncompliant behavior. The factors, it was assumed, did have discoverable

interrelationships that would ultimately be predicted in a theory (Ellwood, 1924).

The development of theory identifying the relationships among the multiple factors associated with illegal conduct was not easy, and the resulting theories within sociology are somewhat vague and general. Even so, tax compliance researchers have not made use of the theory that is available in the sociology of deviant behavior. Three theories which explain some of the interrelationships are Edwin Sutherland's theory of differential association, the theory of culture conflict, and Robert Merton's structural theory of deviance.

Edwin Sutherland's treatise on criminology first emphasized the theory of differential association in the 1939 edition (Quinney, 1975, p. 8). The general idea was that criminal behavior was learned behavior in the very same way noncriminal conduct was learned. Thus young people exposed to adults and peers who acted criminally, learned the techniques, values, and attitudes of illegal behavior. The multiple factors remained, but the causal impact of the factors depended on the "differential association" an individual experienced between factors producing criminal conduct and those yielding law-abiding conduct. Sutherland's socialization approach, while still vague, allowed for the fact that the causal impact of factors might vary from one individual to the next (Sutherland and Cressey, 1970).

Thorsten Sellin applied the culture conflict theory to criminal behavior in Culture Conflict and Crime (1938). The concept of culture conflict was one of the early triumphs of American sociology. In 1918 W.I. Thomas and Florian Znaniecki prepared a multi-volume study which used correspondence between recent immigrants to the United States and others, mostly relatives, who remained in the homeland. The study, called The Polish Peasant, dramatized the differences between the value system of the Polish peasant in Poland and the values which were needed in the industrial, urban United States. Many instances of running afoul of laws seemed obvious cases of conflict between two cultures. Just to give an example, mandatory school attendance, a reform of the Progressive Era intended to reduce child labor, conflicted with the cultural tradition of commanding the labor of one's children. The criminal code and criminal justice system represented the norms of one culture which were imposed on citizens who did not share the same cultural values.

Robert Merton developed a structural model to explain noncompliant behavior (1938). He argued that social structure consisted both of goals and of acceptable means of attaining those goals. The means of goal attainment were not necessarily the most expedient means, but instead were restricted to some subset of all possible means. Thus even given a consensus about goals, the regulatory norms for realizing those goals might impose quite different barriers

to achievement for people of different social positions. Deviance might sometimes be explained as the result of violations of regulatory norms where the individual pursues a normal aspiration in a context in which the regulatory norms are a barrier for that individual's achievement.

Neither the culture conflict nor the social structural theories have been used to deduce testable hypotheses of tax compliance. In fact, tax compliance research is probably best characterized as still reflecting a multiple factor approach. One excellent survey of tax compliance research was even organized around the factors various researchers had selected for inclusion in their studies (Jackson and Milliron, 1986). Unfortunately, there is a strong tendency in empirical work in tax compliance to commit the absurdity which the multiple factor explanation was designed to negate. Specifically, factors are studied as if each has independent causal power. Hite's "The Effect of Peer Reporting Behavior on Taxpayer Compliance" (1988) illustrates the state of the use of sociological theory in tax compliance research.

In Hite's experiment, type of income and the reporting decision of a peer are manipulated. Hite created a manipulation in which some subjects were told a peer was compliant, while others were told a peer was noncompliant. Then she tested for a significant effect of this manipulation on compliance. Hite treats a single peer reference as though it could be isolated as an independent causal factor in tax

compliance. Hite cites "referent theory" and discusses various studies which have involved measures of association between knowing noncompliant taxpayers and tax compliance (or attitudes about tax compliance). She then concludes, "Research in taxpayer noncompliance has begun to recognize this behavioral factor" (p. 48). In general, the question is whether any "behavioral factor" can be treated as an independent cause or not. More specifically, the question is would any advocate of referent theory predict that one peer reference would impact an individual's compliance level?

Suppose, in contrast to Hite's approach, that referent theory had been placed within Sutherland's framework. A survey approach using the differential association framework is possible. In a survey one might identify the influential peers of taxpayers and collect information on taxpayers' knowledge of the compliance behavior of those peers. Experimental approaches might also be improved using Sutherland's framework. A within-subject design might measure the change in compliance by an individual who learns of a new referent's choice. Repeated measures of the effect of the attitudes of referents on the compliance choices of the subject would improve the chances of detecting the effect predicted using referent theory. Sutherland's theory of differential association, or even the multiple factor theory, is persuasive that no single factor can be meaningful when treated as though it is an independent cause of compliance.

Culture conflict or "anomie" may provide more meaningful insights into geographical differences in compliance than have so far emerged.

The tax research based on theories from the general social sciences uses measures of the association between compliance and social and psychological variables which are thought to be related to obedient and to unlawful conduct (Weigel, Hessing and Elffers, 1987; and Smith and Kinsey, 1987). The research question in this work is: Given an amount of taxable income, how much will the taxpayer report? The explanatory variables of interest in this research traditionally consist of items like age, sex, past experience, perceived norms of peer groups, and so forth (Jackson and Milliron, 1986). Economic factors such as opportunity and expected benefit are also considered to be important (Spicer, 1986). Ethical constructs such as victimization from fiscal inequity have been studied (Spicer, 1980). Almost unclassifiable constructs such as the perceived complexity of the tax law have also been examined with an eye on how complexity might change compliance (Milliron, 1985a and 1985b; and Long and Swingen, 1987).

Given the known shortcomings of the factor-by-factor approach, it should not be surprising that this line of research has produced ambiguous and confusing findings. This research includes no studies which include payments or refunds as variables, and therefore theories of obedience

offer no guidance regarding PR effects. It is probably fair to say however, that the reliance on social and psychological factors exogenous to the tax filing logically implies that PR effects are thought to be insignificant.

Theoretical Background: The Economics Tradition

The rational choice model used in economics rests on some virtually indisputable assumptions, such as that people prefer more to less, that goods are scarce, and that an option a person chooses is preferred to an available option not chosen. These assumptions have been used to derive the classical model of markets. The model was kept simple by very restrictive, and indisputably false, assumptions about the information typically available in economic decisions. For example, the classical model employs the assumption that all relevant information is known by all parties to a transaction and that all outcomes are known with certainty. However, the insights from the classical models were extended to situations in which outcomes were not known with certainty and to situations in which parties to a transaction did not have the same information. In accounting research, these two extensions are commonly termed decision making under uncertainty and the agency model, and both extensions make the insights of economic theory applicable to many nonmarket transactions. One of these new applications is to taxpayers' filing and payment decisions.

Utility maximization is a truism under certainty and complete information. That is, what a person chooses is preferred over unchosen alternative options. The notion of utility maximization was extended to decisions under uncertainty in a model called expected utility theory. Expected utility theory was codified as a set of axioms by 1944 by Von Neumann and Morgenstern. During the 1960s and 1970s the properties of this model were extensively investigated by some of the most brilliant economists of the generation such as Arrow (1971), Friedman (1948), and Pratt (1964). Expected utility theory was explicitly applied to illegal conduct by Gary Becker (1968).

The extension of economic theory to transactions in which the parties did not have the same information was accomplished in what accounting researchers know as the agency model (Mirrlees, 1976). The underlying model is essentially one of asymmetric information in which the taxpayer has private information about his income. The taxpayer also has an economic incentive to understate taxable income to the government because the taxpayer retains unpaid taxes. The government investigates filings and imposes penalties for discovered noncompliance. The variables determining compliance, according to this model, are the benefits and penalties of noncompliance and the chance that the penalties will be imposed. The taxpayer weighs the benefit of retaining money legally owed in taxes against the

risk and cost of penalties, and complies only to the extent that the expected costs of noncompliance outweigh the benefits. The justification for the model is that rationality in choice implies that the option selected must be preferred to the options not selected. Since the model describes choices in terms of the objects chosen, it is sterile without a theory of valuation which explains how costs and benefits are valued.

According to the economic theory of tax compliance, expected utility is the valuation model through which the taxpayer performs the relative weighing of costs and benefits. The economic theory of compliance based on asymmetric information and expected utility theory was developed by the early 1970s (Allingham and Sandmo, 1972). The basic model is still used in current analytical studies of compliance (Reinganum and Wilde, 1985, 1988). The risk of detection and penalties have been found to be significantly related to compliance in many empirical studies (Witte and Woodbury, 1985; and Madeo, Schepanski and Uecker, 1987). The research question addressed in this research is: Given a detection risk and penalty schedule, how much taxable income will the taxpayer report?

The timing of payments is not a variable of interest in this research. The timing of payments could only effect compliance through the valuation of benefits, risks, or costs. Absent a time value of money, the timing of payments

has no effect on the monetary value of benefits and costs. Since risk preference in expected utility theory is treated as exogenous, the timing of payments also has no role in the valuation judgment.

Some New Theories Of Rational Choice

Students of decision making under uncertainty in economics, consumer behavior, and psychology found that empirical observations did not always square with expected utility theory. They made use of common violations of some of the axioms of expected utility theory to justify alternative models of rational decision making under uncertainty. These newer models of rational choice, in contrast to the traditional theories of tax compliance, offer a source for a theory which can be used to predict a PR effect.

The common feature of the newer theories of valuation is the realization that choices are not wholly objective. That is, the options valued in decision making do not have features which are invariant to an observer's presence. This makes judgment under uncertainty subjective in an entirely new way. It means that the options in any given choice are defined in the process of being perceived. The effect of the act of knowing on what is known is an increasingly important theme in epistemology in general and philosophy of science specifically (Polanyi, 1974). In the Allais Paradox, most people reverse their preference ordering for two identical

options in two different choices when a third option is introduced in one of the choices. The relative values of the two options being compared do not depend on objective features of the options, but instead vary with the setting within which the options are presented (Machina, 1987; and Kahneman and Tversky, 1979).

Various researchers have emphasized different aspects of the problem. Among economists, Machina has been a leading critic of expected utility theory. His emphasis has been that utility is not linear in risk (1983, 1987). Others have stressed the framing of options. In consumer behavior, Thaler argued that a system of mental accounting is used in which options are valued variously in utility depending on the mental account in which they are registered (1985, 1980). One very notable synthesis of the choice definition and valuation process is prospect theory. In psychology, Kahneman and Tversky have developed an editing and a valuation model as an alternative to expected utility theory. Prospect theory, because it encompasses both the editing of a choice and the valuation of options, is the newer theory utilized in this study. Using prospect theory, a variable previously considered to be irrelevant to the decision, namely the timing of the payments, may turn out to be influential. The timing of payments may produce variation in the way the choice is defined. That variation, in turn, can be predicted to produce a PR effect.

Prospect theory was developed by Kahneman and Tversky (1979) to account for apparent anomalies in expected utility theory. Kahneman and Tversky's most striking claim is that people are risk-averse with respect to gains and risk-seeking with respect to losses. Their evidence for this view is based on subjects' stated preferences for prospects (risky outcomes). For example, they found that most people state that they prefer a certain gain of $3,000 to an 80% chance of winning $4,000, a risk-averse choice since the expected value of an 80% chance of winning $4,000 is $3,200. If the prospect is negative, on the other hand, most people state that they prefer a risk-seeking option. That is, most people state that they prefer an 80% chance of losing $4,000 to a certain loss of $3,000. Of course the expected value of the 80% chance of losing $4,000 is a loss of $3,200. So the amount of risk people are willing to be exposed to varies depending on whether the prospect is a gain or a loss. This risk reversal can be shown to be inconsistent with expected utility theory.

The risk preference utilized by the decision maker depends on whether he has classified the prospect as a gain or as a loss. According to Kahneman and Tversky, an initial step in decision making under uncertainty is an editing process, one component of which is defining the prospect as a gain or as a loss. The perception of the uncertain outcome as a gain or as a loss is a result of the frame of reference

the individual adopts for visualizing the options. This means that identical options could be viewed as a gain by an individual using one frame of reference and as a loss by an individual using a different frame of reference. The example of the loss prospect given above involving an 80% chance of losing $4,000 might be viewed positively as a 20% chance of gaining $3,000.

Kahneman and Tversky believe that the current asset position is the frame of reference most commonly used to classify options as gains or losses. The current asset position involves the amount at issue in the immediate choice. For example, if an individual faces a prospect with outcomes of either losing $10 or losing $3, then the choices would generally be valued as minus $10 and minus $3. Since the outcomes are negative, an individual facing such a prospect would be willing to accept more risk than he would if the same amounts were at stake as gains.

In expected utility theory, the presumption is that all outcomes are valued in terms of the ending wealth position. Thus the outcomes of the same losing prospect for a person with wealth of $5,000 would be valued as plus $4,990 and plus $4,997. If the ending wealth position is the frame of reference, both outcomes are still positive numbers. Though Kahneman and Tversky have a realm for losses, in expected utility theory losing lotteries are mere declines in a positive wealth balance. To use accounting jargon, expected

utility theory assumes a balance sheet view of uncertain outcomes, while prospect theory assumes an income statement view of uncertain outcomes.

The frame of reference influences the valuation not only by fixing the sign of the outcomes but also by determining the magnitude of the values in the prospects. The same dollar difference between two options in two different choices could be valued differently. For example, saving $5 on a $15 calculator would likely be valued more highly than saving $5 on a $300 television. Although in each setting, $5 is at stake, the $5 will be valued differently if the frame of reference is the purchase price. Specifically the relevant values are $10 and $15 for the calculator and $295 and $300 for the television set. The value function would typically be steeper between $10 and $15 than between $295 and $300 (Kahneman and Tversky, 1982, p. 168). If the ending wealth position were the frame of reference for these two options, then saving the $5 should have virtually the same value in both purchases.

The application to individual income taxes involves the frames of reference available in the tax context. There are several potential frames of reference a taxpayer may adopt. Most of the plausible frames of reference are either consistently positive or consistently negative. One frame is the ending wealth position after paying taxes, which is the frame of reference anticipated under expected utility theory.

This frame of reference would yield a positive sign to the outcome for all taxpayers who are not bankrupt, and, absent a time value for money, the timing of the payments would be irrelevant in valuing tax outcomes. Other plausible frames of reference include the aftertax income, the tax liability, and the settlement amount. The aftertax income should always yield a positive sign for tax outcomes. The tax liability, if adopted as the frame of reference, would almost always produce negative outcomes. The settlement amount, however, could yield an outcome which is either positive or negative depending on the level of prepayment. The settlement amount appears to be the more salient value for most taxpayers (Westat, 1980, p. 11). According to prospect theory, use of the settlement amount as the frame of reference may result in valuation of payment-due and refund tax filings under different risk preferences.

Prospect theory can be substituted for expected utility theory as the valuation model used to evaluate risky choices while the agency model is retained. If prospect theory applies, there is a potential for payments and refunds to change the frame of reference used by a taxpayer in valuing tax filing options. Specifically, if payment-due filings are framed as losses and refund filings are framed as gains, then taxpayers may make risk-seeking filing choices in payment-due situations and risk-averse filing choices in overpayment situations. Prospect theory, therefore, is a theory of

valuation which offers some predictions which conform with the beliefs of many tax practitioners that filing with a payment due versus filing for a refund alters the behavior of taxpayers.

Hypotheses

The mainstream theories used in tax compliance research lead to the conclusion that a PR effect does not exist. Prospect theory leads to the conclusion that a PR effect on compliance may exist. Anecdotal evidence from tax practitioners also supports the notion that a PR effect is real. The research question is: Is there a PR effect on taxpayer compliance? The hypothesis related to this research question is stated in the alternative form.

Ha_1: There will be a significant PR effect on the amounts reported in tax filings.

According to prospect theory, the PR effect identified in the first hypothesis is expected because subjects are expected to prefer riskier filing choices in the payment-due situations than in the refund situations. Because the amount reported and the risk are related to each other, the same hypothesis can be stated in terms of the level of risk chosen in refund and payment-due situations. The second research question is: Is there a PR effect on the level of risk preferred in tax filings? The hypothesis related to this research question is stated in the alternative form.

Ha_2: There will be significant PR effect on the level of risk preferred.

A third hypothesis has to do with the influence of context. Though evidence of the changes in risk preferences predicted in prospect theory have been found in many contexts, only the most limited evidence of such an effect has been found in tax filing judgments. A context contrast, therefore, may provide insight into differences between taxpaying decisions and decisions in other contexts. A second motivation for using context comparisons comes from the use of cash incentives in the experiment. The context comparisons are a control check against distortions of the responses due to the cash incentives. Should the cash incentives induce disingenuous gaming behavior, then the context contrasts will fail. The research question stemming from these two considerations is: Does context affect the willingness to be exposed to risk? The main hypothesis related to this research question is stated in the alternative form.

Ha_3: There will be a significant effect of context on the level of risk preferred.

These hypotheses are stated as test hypotheses in Chapter 4. The specific tests of the hypotheses depend on the details of the research design, which is also the topic of Chapter 4.

CHAPTER 3

REVIEW OF RELEVANT RESEARCH

Introduction

The research relevant to a PR effect comes from diverse sources. Some of the research is motivated by contradictions in expected utility theory. This research includes the work of Kahneman and Tversky, which has been extended to medical and public policy settings. One theme in this work is discussed, namely the ambivalence over whether a change in risk preference is essential to utility maximization or is a bias. The applications of prospect theory in tax research are also discussed. Tax research using prospect theory as a framework to test for a PR effect has produced only the most limited evidence of a PR effect.

Risk Preference Reversal: Utility Maximization Or Bias

Prospect theory merged the well-known tendency for people to be risk-seeking in loss situations yet risk-averse in gain situations and the formal critique of expected utility theory (Kahneman and Tversky, 1979, p. 268; and Machina, 1983, p. 57). Risk preference reversal between gain and loss situations results in preferences inconsistent with

the ordinal rankings predicted in expected utility theory. Therefore, instances of inconsistent preferences stemming from risk preference reversal were collected to create a telling critique of expected utility theory. The process of documenting instances of preferences inconsistent with expected utility theory was, at the same time, a process of documenting the ubiquity of risk preference reversal in decisions under uncertainty.

The linking of risk preference reversal with cases of inconsistent preferences resulted in an enormous ambivalence which pervades prospect theory. It appears that the common preference for riskier choices in losses than in gains would imply that people maximize utility (or the value function of Kahneman and Tversky) through risk preference shifts. The ambivalence is that risk preference changes are associated with inconsistent preferences. With changing risk preferences, pairs of choices can be devised such that option A is preferred to option B in one choice, but option B will be preferred to option A in the other choice. Since only one option can be most preferred, it seems that the inconsistent preferences can not result in outcomes of greater value than if a single preferred option was consistently selected in both cases. The question is whether risk preference reversals, which are endemic, result in less value than would exist if they did not occur.

Prospect theory is concerned both with how perceiving a choice influences the definition of the options and with the valuation of the options once perceived. Kahneman and Tversky appropriately call the perception phase editing. Some of the editing aspects of judgment appeared in Tversky's judgment model called elimination by aspects (Tversky, 1972). According to Kahneman and Tversky, preference reversal may result independently from several of the editing procedures. Preference reversal can occur from a sequential formulation of problems as compared to formulations based on probabilities of final states (1979, p. 271). Preference reversal can also occur due to the order in which simplifications are performed (1979, p. 275). Preference reversal can also occur due to the certainty effect as in the Allais Paradox (1979, p.266). The preference reversals from sequential formulations of problems, the Allais paradox, and the order of simplifications may produce a choice which is different from the choice the decision maker would have made if he had been able to process the data precisely and simultaneously. They are associated with economizing cognitive resources and therefore with bounded rationality.

The sources of preference reversal associated with bounded rationality were not the only sources for inconsistent preferences. The remaining source was risk preference shifts. Kahneman and Tversky argued that the frame of reference assumed in expected utility theory, the final

wealth position, was not the frame of reference people actually used. They claimed that the amounts at stake in a choice, not final wealth positions, were the outcomes valued. Thus a risk of a loss was valued in a loss realm. The frame of reference, gain or loss, controlled risk preferences.

Kahneman and Tversky documented the pervasive influence of the frame of reference showing that risk-seeking choices in losses and risk-averse choices in gains was generic. But the case for a generic risk preference reversal was established through manipulations to prompt subjects to adopt one frame of reference rather than another. They also showed that the frame of reference could be induced by the manner in which a choice was presented.

From this perspective, risk preference shifts were just like the other editing steps; they were caused by frames of reference which individuals could no more control than they could simultaneously process multi-step lotteries. Could there be any more ironic outcome from this critique of expected utility theory? It produced the conclusion that though risk preference shifts are generic to human preferences, they are also a bias in decision making under uncertainty. Ending wealth position, the frame of reference assumed in expected utility theory, was the frame of reference recommended by Kahneman and Tversky to minimize the bias from framing (1979, p. 287)!

The ambivalence toward risk preference shifts present in their 1979 formulation of prospect theory is present in subsequent research. Risk preference reversals induced by framing manipulations are often treated as an explicit bias in judgment (Harwood, Pate, and Schneider, 1988), but risk preference reversal as a generic quality of human preferences is still of interest as the source for particular institutional arrangements (Thaler, 1985). The general pattern seems to be that when the risk preference changes are accomplished by framing manipulations, and therefore have a source external to the decision maker, then reversal is considered a bias. When institutional arrangements seem to be inherently gains and losses, the risk preference shifts are not evaluated as though they were a bias. Some extensions of expected utility theory in the Pratt-Arrow tradition exemplify an acceptance of unbiased preference reversals. Some researchers hypothesize an s-shaped utility function (as do Kahneman and Tversky), but the frame of reference is always the ending wealth position. All outcomes are always positive. An s-shaped utility function implies risk preference reversal. The idea is that individuals are both risk-averse and risk-seeking depending on their wealth level and the proportion that the amount at risk is to their wealth level. In this line of thought, there seems to be no hint that risk preference changes are a bias (Machina, 1983, pp. 53-56).

The ambivalence toward risk preference shifts means there are conflicting views of such shifts. The research design to study the PR effect as well as the policy implications of the PR effect depend on whether one believes the risk preference changes signify a bias in judgment or value maximization. For example, an interest in risk preference reversal as a bias suggests framing contrasts to show that the choices of subjects can be manipulated. From a policy perspective, alternative frames of reference might be evaluated in terms of the effects on compliance, public revenue, or taxpayer utility. The question of whether risk preference shifts constitute a bias or value maximization is kept open in this study by considering whether or not taxpayers prefer a particular frame of reference for making tax filing judgments.

Applications Of Prospect Theory

Kahneman and Tversky used very general hypothetical gambles with students as subjects. For example, a subject might choose between an 80% chance of $4,000 and a 100% chance of $3,000. The risk preference reversal was implemented by stating the same choice in negative amounts (choose between an 80% chance of minus $4,000 and a 100% chance of minus $3,000). The relative slopes of the value function were traced with similar choices. They presented their findings primarily in terms of dollar lotteries, and

they warned that variation was sure to occur when specific parameter values were used (1979, p. 278).

Nonmonetary situations were created to see whether or not the risk preference reversals over hypothetical lotteries would also occur in nonmonetary situations. Kahneman and Tversky and other researchers created nonmonetary gambles such as choosing from alternative programs in response to an outbreak of disease. When the options included an alternative which was stated in terms of number of people saved, risk-averse choices were most common. When the options included an alternative which was stated in terms of number of people dead, risk-seeking choices were most common. Scenarios involving medical treatment produced the same risk preference reversal in response to manipulations of the frame of reference (Kahneman, 1986 p. S254). Scenarios involving many additional contexts have been developed, and framing manipulations using anticipated events and social norms produced preference reversals (Kahneman and Tversky, 1981, 1982; Kahneman, 1986; Thaler, 1985; and Kahneman, Knetsch and Thaler, 1986).

The risk preference reversals have also been studied in business contexts. Many investment professionals believe that investors tend to hold losers too long and sell winners too early. Shefrin and Statman applied prospect theory to investments in securities to develop a testable theory (1985). Harwood, Pate and Schneider tested for risk

preference reversal in business decisions, using as subjects MBA students many of whom had business experience. Their idea was to determine whether or not risk preference reversal would occur in a context in which accounting information was typically used by expert decision makers. They found preference reversals, which they interpreted to be a bias (1988). There have been other applications of prospect theory to business decisions (Levin, Johnson and Davis, 1987). One of these applied prospect theory to multi-attribute business investment outcomes (Payne, Laughhunn and Crum, 1984). Since choices over multi-attribute options have potential for producing preference cycles, applications of prospect theory to the multi-attribute case is very important (Machina, 1983, pp. 36-38). As was true of the single attribute business choices, Payne, Laughhunn and Crum found preference reversals in multi-attribute prospects.

There are very few studies in the tax compliance area that are concerned with any predictions from prospect theory. One of these has to do with what Kahneman and Tversky call the certainty effect, which is the tendency to overweight risks close to zero or one (Jackson and Jones, 1985). Jackson and Jones did find evidence of the certainty effect, but the study was not designed to capture the risk preference reversal phenomenon. The other tax compliance research using prospect theory involves a search for evidence of PR effects. The results of the current research taken as a whole is

ambiguous. Two studies found a PR effect on compliance conditional on gain or loss framing exogenous to the experiment (Chang, Nichols and Schultz, 1987; and Robben, Webley, Elffers and Hessing, 1988). Another found a PR effect only when refunds were explicitly labelled to be gains (Schadewald, 1989). Another found PR effects only for low-income subjects (Hite, Jackson and Spicer, 1988). These behavioral studies have not produced convincing evidence that PR effects are real.

Chang, Nichols, and Schultz (1987) used a within-subject design with six cases. The amount of savings, risk of audit, and penalty rate were manipulated, while the dependent variable (compliance) was dichotomous. The subjects were 56 MBA students who made generally risk-averse choices. In debriefing, the subjects stated whether they viewed tax payments as reduced gains or as sure losses. Chang, Nichols, and Shultz divided the subjects into two groups assuming that subjects who claimed that tax payments were sure losses had framed the compliance choice as one among negative payoff lotteries and that subjects who claimed that taxes were reductions of income had framed the cases as positive payoff lotteries. By splitting the subject group this way, they found significantly more noncompliant (risk-seeking) choices among those stating that taxes were losses than among those who claimed taxes were reductions in income.

A study by Robben, Webley, Elffers and Hessing (1988) reached very similar results. They embedded two years of tax filings in a general business simulation case. They manipulated the withholding that was reported for the year to create payment and refund settings. Like Chang, Nichols and Schultz, they found generally risk-averse choices in both payment-due and refund cases, and they found no significant PR effect. In debriefing, they asked the student subjects to recall whether the tax outcome had been a payment due or a refund. Those who recalled that the tax outcome had been a payment due preferred riskier filings than did those who were unable to recall the sign of the settlement amount.

Schepanski and Kelsey (1988) created three tax filing settings: a refund setting, a loss setting, and a final asset setting. The compliance options in the three settings had identical expected values. The loss filing was an unexpected additional payment due. The refund setting involved an unexpected reduction in an unexpected refund. The final asset setting stated the unexpected reduction in terms of final wealth positions instead of in terms of either a payment or a refund. Subjects were assigned to one of the framing conditions. Within each framing condition, repeated measures were taken over five levels of audit risk and penalty. The subjects made significantly riskier choices (lower compliance) in the loss framing than in the other two conditions and significantly less risky choices (higher

compliance) in the asset position framing than in the other two. In this study, the effect of the frame of reference is considered to be a bias.

Hite, Jackson and Spicer (1988) manipulated the amount of the total payment or refund amount while keeping constant the expected value of noncompliance. They found no significant effect from the amount or the sign of the settlement amount for the juror subjects in general, but did find some PR effect on noncompliance for the low income subjects.

Schadewald (1989) was interested in discovering whether the frame of reference for subjects' framing of gains versus losses was the total liability as compared to the payment or refund. His cases involved an explanation to subjects of an expected outcome followed by a realization, which meant that the cases included a variance of the liability from expectation. Of course, he also had the payment versus refund at settlement as a potential reference point. He found no PR effects and no effects of the variances on the subjects' preferences for aggressive tax return positions. He administered one version of the instrument in which the outcomes were labelled as gains or as losses. He found that the subjects' choices were more risk-averse in the cases labelled gains than in the same cases with no labelling.

Limitations Of The Existing Tax Research

Choosing an appropriate research design is a problem in any study of a PR effect. The problem is that a PR effect involves personal preferences which are idiosyncratic aspects of rational choice. To use analytical approaches and optimization approaches one must assume a utility function which produces valuations. The results obtained through analytical studies depend on the utility function selected. For most such work, an assumed valuation model such as a Von Neumann-Morgenstern risk-neutral utility function provides a plausible benchmark allowing identification of the relationships among the variables of interest in the model. The utility function itself is generally not one of the variables of interest. But, if risk preferences are endogenous to the problem under study, then the utility function is the object of study. Consequently, an empirical approach rather than an analytical approach is needed.

Econometric techniques fail for lack of data. Taxpayers either faces a payment due or a refund. The hypothesis of a difference from the two outcomes is testable with real filings only to the extent that subjects in the two conditions are comparable. A longitudinal approach requires the collection of multiple observations on a taxpayer in which the filings differ on payments and refunds. Flaws in this approach include the difficulty of obtaining returns with variation confined to payment level and the likelihood

that individuals' utility functions change over time. A cross-sectional approach has even more severe difficulties in controlling for the confounding variables which may influence compliance than does the longitudinal approach. Econometric approaches promise insurmountable data and modelling problems.

Experimental methods, therefore, are an appropriate choice. However, the current research has important limitations. First, there are no incentives aligned with the outcomes of the subjects' choices. Subjects who choose risky outcomes and subjects who choose certain outcomes all receive the same payoff. The theory utilized in the studies is that the subjects value monetary outcomes in utility and that risk preferences are endogenously affected by gain and loss prospects. In these experiments, however, the subjects face no risk, no cost, and no benefit from the choices they make (Davis and Swenson, 1988). Subjects are not asked to report their true past tax filing behavior. Instead, subjects are given a hypothetical scenario and then asked which filing option they would choose. Kahneman and Tversky appear to have done the same in asking student subjects to choose preferred lotteries. Kahneman and Tversky argue that their subjects are able to respond to the hypothetical cases in ways which reflect preferences over real world lotteries (1979, p. 265). Kahneman and Tversky's lotteries involve simpler situations than preparing a tax filing. Kahneman and

Tversky's lotteries are gambles against nature with no moral hazard and often only two choices. By contrast, tax filing judgments involve gambles with an interactive government and multiple options. The task in applying prospect theory to tax filings is to retrieve preferences across tax filing options, an area in which self-insight into judgments may be very low. Taxpayers may not be able to reliably report past compliance behavior (Elffers, Weigel and Hessing, 1987), much less estimate probable choices in hypothetical contexts.

A second limitation is the use of student subjects (the use of jurors in the Hite, Jackson, and Spicer study is an interesting exception). Even when the students used are experienced filers, conducting the experiment in a business class meeting or as a business course exercise may influence the subjects' choices. It would appear to be much better to use experienced tax filers. A third limitation of the current research is the motivation of payments and refunds as variances in the liability rather than as differences in the timing of the tax payments. If the payment due or refund is merely a positive or negative variance in the amount of the liability, then any risk preference reversal from the PR effect is confounded with any risk preference reversal from the variance. Put in Kahneman and Tversky's terms, if the frame of reference is the expected payment due or refund, then outcomes poorer than expectation may be treated as losses and outcomes better than expectation may be treated

as gains. Such a frame does not necessarily coincide with the settlment amount as a frame. For example, one may get a refund which is less than was expected. It is unclear in such confounded designs whether the variance is intended to be available as a potential frame of reference or whether it is introduced to make the payment or refund plausible. Commonly the payment or refund is positively correlated with the variance (Schadewald, 1989; and Schepanski and Kelsey, 1988), so that, for example, a negative variance is associated with a payment due.

New Directions For Research On The PR Effect

Current tax research using prospect theory contains no convincing evidence of a PR effect in tax filing judgments. This is puzzling because prospect theory seems applicable. Risk preference reversal is not thought to be specific to a context, and tax filings are thought to be decisions under risk. Risk and penalties have been consistently found to affect tax compliance. Tax filings seem to involve a salient parameter, the settlement amount, which varies in sign and therefore ought to be used as a frame of reference which would result in risk preference shifts. The summary finding of the research so far is that tax filing choices are uniformly risk-averse and that the sign of the settlement amount has no impact on the amount of the tax liability reported.

The absence of an effect commonly elicited in other settings means that an understanding of the PR effect in tax filing settings has potential for adding to our understanding of taxpayer judgments. One logical conclusion from the current research is just that; people make tax filing decisions differently from the way they make other risky choices. Another possible explanation is that the experimental manipulations were faulty or were too weak. Therefore the virtual absence of significant PR effect in prior research is not unambiguously interpretable as a finding that tax filing judgments do differ from risky choices in other contexts. Even if we did conclude that the tax filing judgments are different, we still have no clue as to the features of the tax context that account for the differences. By clarifying differences between the tax filing judgments and other risky choices, continued research of the PR effect can give us new insight into taxpayer behavior.

The limitations of the existing research provides a basis for innovations in the research design used in this study. The first innovation is to eliminate an important confounding variable. In several prior studies of the PR effect, the variance of the tax liability from an expected liability was used to motivate variation in settlement amounts. That is, a payment due is described to be the result of a tax liability higher than expected, and a refund to be

the result of a tax liability lower than expected. This confounds the effect of the settlement amount with the effect of the variance. Since there are no variances in this study, variations in the settlement amount is produced solely by differences in prepayments. The elimination of this confound should make it less likely that a PR effect would be detected. Three other innovations, however, should strengthen the tests. These three innovations are use of cash incentives in the experiment, the use of a within-person design, and use of context manipulations. These innovations allow more powerful tests of the PR effect.

CHAPTER 4

THE RESEARCH DESIGN

Introduction

A formal description of the design opens this chapter. It covers the variables, the model, the hypotheses and the statistics for testing the hypotheses. The experimental procedures are described next. These procedures include the nuts and bolts of case construction, recruitment of subjects, and the description of the experimental sessions. The design innovations, particularly the reasons for the use of cash incentives in the design, are described next. The design innovations are followed by the manipulation and control checks. The statistics and measures for these checks are identified and explained. The final topic is the implication of the empirical results for utility representations of preference orderings.

Variables And The Model

There are four variables in the general model.

COM = "Compliance"

R = Risk

P = Payment or Refund level

CON = Context

Compliance (COM) is the percentage of an item at issue included in taxable income when the correct amount to include is uncertain. Note that COM is not the percentage of a certain, legally correct amount reported, the usual concept of tax compliance, and COM does not measure illegal choices. Here COM applies to an item at issue for which the correct amount to report is uncertain. This formulation is compatible with over-compliance, and in the experimental task, overstating the tax liability was more likely than understating the tax liability.

Risk (R) is the probability of the occurrence of the unfavorable outcome. For the tax cases, R is the chance that the option selected understates the subsequently determined correct amount to report for the item at issue.

Payment (P) is the amount of the settlement. This amount does not include any payments from a subsequent finding of understatement. There are three levels of P: refund, low payment, and high payment.

Context (CON) is the context of the experimental tasks. There are three levels of CON. They are making a tax filing, purchasing health insurance, and selecting a gamble. The nontax contexts were selected because risk preferences in these two settings are well known and extreme. One of the conundrums of utility theory is how the same person can buy insurance and gamble.

COM and R are dependent or response variables, and CON and P are manipulated, independent variables. The model, in general form, is

$$\text{COM, R} = u + \text{CON} + \text{P} + \text{error}.$$

With the two independent variables each having three levels, a full model would be a 3 x 3 design with nine cells. Some of those nine cells are not meaningful. Additionally, evaluating the lotteries was effortful so that a minimum number of cases needed to make the experimental contrasts was chosen. There are five cases: three tax cases, one gamble, and one health insurance case. The design is illustrated in Figure 4-1.

The underlying variables for COM and R are continuous variables, but each is measured or manipulated using a discrete number of possible values. COM can take on values of 100%, 75%, 50%, 25% and 0%. R can take on values of 0%, 15%, 25%, 33% and 40%. The five available levels of COM and R are treated as though they are responses from a continuous distribution. In pilot testing, subjects made a choice from among four possible levels of COM (100%, 67%, 33% and 0%). They were also asked to indicate a preferred COM (anywhere between 0% and 100%). Few subjects selected a continuous value different from the available choices. Therefore, five levels were selected under the assumption that these five levels allow for responses which are like those that would be selected from a continuous distribution.

		CONTEXTS (CON)		
		Tax Filing	Health Insurance	Gamble
PAYMENT LEVELS (P)	Refund	X		
	Low Pay	X		X
	High Pay	X	X	

Figure 4-1
The Experimental Design

P is also a variable for which the underlying values are continuous, but P is used as a categorical variable with three levels which are refund, low payment, and high payment. P was not treated as a continuous variable because the exact dollar value of P varies within each case depending on the subject's choice of COM and R. In the low payment case, P has dollar values from minus $200 to plus $200 depending on the option selected in the low payment case by the subject. In like manner, the dollar value of P in the refund case varies from plus $400 to plus $800 depending on the reporting option selected by the subject, and the dollar value of P in the high payment case varies from minus $400 to minus $800 depending on the reporting option selected by the subject. Therefore, although the underlying variable is continuous, P is treated as a categorical variable with three levels. The refund case is unambiguously positive. Even with an unfavorable outcome, the net settlement remains positive. The high payment level is unambiguously negative; even the most favorable outcome yields a net payment. The level called low payment, however, has parameter values of a mixed lottery.

The final variable is used to implement the context manipulation. The contexts are qualitatively different. The context variable, CON, is therefore a categorical variable. It has three levels: make a tax filing, purchase health insurance, and choose a gamble. The health insurance case has the dollar values of the high payment tax case and the gamble

case has the values of the low payment tax case. The risk, R, in the health insurance case is a risk of needing an uncovered medical service, and the risk in the gamble is the risk of getting the unfavorable outcome (losing the bet). COM still has an interpretation in the health insurance case. It is the percentage of the cost of the full coverage policy paid for a cheaper policy which covers fewer services. COM does not have an interpretation in the gamble, so all comparisons across contexts use R, the risk level preferred.

The Hypotheses

Hypothesis 1

The first hypothesis involves the effect of the settlement amount on the amount of an uncertain item reported in the tax filing task.

Ho_1: There is no significant PR effect on the amounts reported in the tax filings.

The first hypothesis is tested by measuring the effect of the payment level (P) on the percentage of the item (COM) the subject chooses to report when the correct amount to report is uncertain. The statistic for testing this hypothesis is the F statistic for significant within-subject differences in the variable COM across the three tax cases. This is a simple one-way analysis of variance with three levels of P and one level of CON. Repeated measures on COM selected in each of the three tax cases yields the within-subject test. The alternative hypothesis is also directional.

That is, it is expected that COM will be higher in the refund case than in the high payment case.

The refund case is an unambiguously positive lottery, the high payment case is unambiguously negative, and the low payment case is a mixed lottery. If the settlement amount is used as a frame of reference, then, according to prospect theory, the low payment case might well bridge the inflection point in the value function. For this reason, the precise relationships between the refund and low payment and between the low payment and high payment cases are not predicted.

Hypothesis 2

Hypothesis 2 also involves the PR effect. This hypothesis has to do with the effect of the settlement amount on the level of risk preferred for the item at issue.

Ho_2: There is no significant PR effect on the level of risk preferred for reporting an item at issue.

The second hypothesis is tested by measuring the effect of the payment level (P) on the amount of risk (R) preferred in the tax cases. Since R and COM are chosen simultaneously in the option selected by the subject, the results should be the same as they were for Hypothesis 1. The statistical test for the effect of P on R exactly parallels that of the test for the effect of P on COM. The test of the hypothesis is an F test for a significant within-subject effect on the R selected by subjects in the three tax cases.

The alternative hypothesis is also directional. That is, it is expected that R will be higher in the payment case than in the refund case. As in the test of Hypothesis 1, the precise relationships between the refund and low payment and between the low payment and high payment cases are not predicted. The precise relationship and direction of differences among the three tax cases is measured with analyses of variance for each of the three pairings of tax cases.

Hypothesis 3

The third hypothesis has to do with the effect of context (CON) on the level of risk (R) of the preferred option. Contexts with well-known, extreme risk preferences were selected to get some reference for evaluating risk preferences in tax decisions. One context comparison is between the purchase of health insurance and the tax filing at the high payment level. The second context comparison is between selecting a gamble and the tax filing at the low payment level. There are two versions of this hypothesis, one for each of the nontax contexts.

Ho_{3A}: There is no significant difference in the level of risk preferred in a tax filing and in a purchase of health insurance.

The statistic for testing this hypothesis is the F statistic for significant within-subject differences in the response variable R in the high payment case in comparison to R in the

health insurance case. This is a simple one-way analysis of variance with one level of P and two levels of CON.

The alternative hypothesis is also directional. It is expected that R will be higher in the payment case than in the health insurance case. The relationship between R in the health insurance case and in each of three tax cases will also be explored. The other part of Hypothesis 3 is the contrast between the tax and the gamble contexts.

Ho_{3B}: There is no significant difference in the level of risk preferred in a tax filing and in a gamble.

The statistic for testing this hypothesis is the F statistic for significant within-subject differences in the response variable R in the low payment case in comparison to R in the gamble. This is a simple one-way analysis of variance with one level of P and two levels of CON. The alternative hypothesis is also directional. It is expected that R will be higher in the gamble case than in the low payment tax case.

Experimental Process And Procedures

Description of the Experiment

The project is a behavioral experiment using cash incentives. Experienced taxpayers are the subjects, each of whom receives a payoff proportionate to the ending balance of an experimental fund of $3,000 which he uses to cover all costs and liabilities in five cases. The fund is reduced by any additional payments due and increased by any refunds. The

fund is converted to cash payoffs at the end of the session at the rate of $.50 in cash for each $100 remaining in the fund balance. The average cash payment was $8.75. One hour was allotted for each session. The sessions were held in various meeting rooms and classrooms at north central Florida area schools and at the University of Florida.

The session begins with a 15 to 20 minute training program. The program describes the three tasks (make tax filings, buy health insurance, and select a gamble), the payoff, the use of the spinner, and the procedures. The training session also includes a sample case, which is a tax case. The resolution of risky outcomes using the spinner is illustrated for each of the five options in the sample case.

Next the subjects receive a package of cases with an attached payment envelope. Upon selecting an option in a case, the subject circles the preferred option. If this option is a filing for a refund, the subject raises his hand, and the refund is given to him. If the option requires a payment, the subject counts out the payment and places it in his payment envelope. The subject then proceeds to the next case. The subjects are not told that there are five cases, and an end page asks subjects to rank the cases simultaneously after looking back over all the cases. Upon completing the simultaneous rating, the subject raises his hand. An assistant then takes a spinner to the subject. The subject makes a separate spin for each case for which a risky

option has been selected. The subject pays the additional cost from any unfavorable spin. After completing all spins, the subject counts up the amount of money he still has. The subject reports the amount and gives the fund to the experimenter. The subject is then given a debriefing questionnaire to complete. Upon turning in the questionnaire, the subject receives his cash payoff.

The training session lasted 20 minutes, the choices in the cases took 20 minutes, and the debriefing questionnaire took less than 10 minutes. The sessions were highly interactive. The maximum number of subjects in any single session was 11. The ratio of assistants to subjects was greater than .33 at every session.

Subjects

Sixty-five experienced taxpayer completed the experiment. Subjects had to meet one of two qualifications to participate. Either the subject had to be at least 30 years old or the subject had to have been fully employed for more than two years. The minimum number of tax filings reported by any participant was four (two subjects), and 78% of the subjects had ten years or more of tax filing experience. Some subjects were recruited through leaflets and others through direct solicitation. The session was described as a study in personal judgments in which the judgments under study were adult spending choices. Paying taxes was given as an example of an adult spending choice.

Subjects were recruited between the middle of June and late July of 1989 with different groups targeted. Fourteen subjects were recruited from parents participating in a Freshman orientation program called Preview at the University of Florida. Some of the participants recruited in these sessions were University staff. Eight subjects were recruited from a local high school athletic boosters group and ten subjects were recruited from an area band boosters group. By prior arrangement, participants from these two different booster groups contributed their cash awards to the organization of which they were members. Twenty-two subjects were recruited from summer school classes in the College of Education at the University of Florida. Eleven subjects were all members of a single church congregation, gathered for a social function.

Experienced subjects were considered essential for two reasons. First, the context manipulation is difficult to interpret unless subjects have some experience with individual income taxes. Failure to detect a context effect would be far from conclusive if subjects were inexperienced. Second, experienced taxpayers were necessary to insure that the subjects understood the timing of tax payments. Earlier designs involving both prepayment and settlement choices proved too complex to administer. In the final design, the experience of the taxpayer was relied on to insure that the relationships among prepayments, liabilities, and settlement

amounts was understood. The regularities reflected in the aggregate data on prepayments suggests rationality. That is, it does not appear that taxpayers visualize their settlement amount to be the product of a stochastic process independent of their prepayments. One study of the rationality of taxpayers with respect to prepayments also supports this view (Moore, Steece, and Swensen, 1985).

Case Construction

The cases were constructed from four criteria. These criteria were that the cases have a specific tax context, that the judgments be decisions under risk, that the choices would carry economic consequences to subjects, and that the cases would yield meaningful data given the experimental design.

The tax filing context is implemented by assigning the subjects the task of completing an individual tax return. The decision remaining to accomplish the task is to choose the amount of an item at issue to report. The taxable income and withholding are given, as are the settlement amounts for various options for the amount of the item at issue to report. The tax cases are general tax filing situations (about 12 lines long) with one unspecified item at issue. Subjects complete three tax cases which differ in the withholding level and, therefore, in the settlement amount. The three cases have withholding levels set at 84% of the

liability, 96% of the liability and 108% of the liability (where 100% of the liability is the no-risk filing option).

To implement decision under risk in a way compatible with the assumptions of expected utility theory and prospect theory required that additional criteria be met. The outcomes had to be uncertain ones. Uncertainty is implemented by presenting the item at issue as one for which the legally correct amount is indeterminable at the time of filing. This can be pictured as though the audit risk and the detection risk are 100%, but there is still uncertainty because the correct amount can be determined only from subsequent events. Subjects were informed that selecting a no-risk option had a high likelihood of overstating the true liability and that none of the choices were illegal choices.

In each case the subject can avoid all risk of understating the liability by reporting all the income or taking none of the deduction. The remaining options in each case carry varying risks of understatement. The taxable income, withholding, tax liability and settlement amounts are all available in each case. The settlement amount, risk and cost of an unfavorable outcome are presented for all options. The options within any case have parameters such that increasingly risky options are intended to be mean-preserving spreads of the certain option. Thus the concept of risk used to create variation in risk is that of Rothschild and Stiglitz (1970).

There are five options in each case to allow subjects to vary their choices of the risk level preferred from one case to the next. The dollar amounts in the cases are rounded to make the dollar and risk differences from one option to another relatively easy to compare by visual inspection. The risk level and cost if the item is understated are identical for all tax cases. To prevent the penalty from decreasing in comparison to less risky options, the risk levels only go up to a 40% chance of an unfavorable outcome in the riskiest option. The rounding of the dollars in the cases yields options which vary in even $100 increments. The rounding results in slight variation from exact, mathematical equality of expected values across the options. In fact, the expected value increases slightly and progressively moving from Option A (no-risk option) to Option E (the 40% risk option).

The dollar parameters are varied in three levels (refund, low payment, and high payment) to create variation in the settlement amount. The dollar values in the options for one level are linear transformations of the values in the corresponding option of the other levels. Thus although the expected value of Option C in the refund case is higher than the expected value of Option C in any other case, the variance and risk in Option C is identical in all cases. Put differently, the risk amount and variance in Option C in any one case are identical to the risk and the variance of Option C in all other cases.

The outcomes from risky choices is resolved using a game-type spinner which implements the risk stated in the option. Thus the outcome of an option with a 25% chance of understatement is resolved by a spin in which landing in the first quarter of the wheel is an unfavorable outcome, and landing in any other part of the wheel is a favorable outcome.

To implement economic choice, cash incentives were introduced. Each subject receives a $3,000 fund of experimental dollars which is used to cover all costs in all cases. Unspent funds are cashed in at the rate of $.50 in cash for each $100 left over in the experimental fund.

The tax versus nontax contrasts involved creation of additional cases. The contexts chosen were two for which risk preferences are well known and extreme, namely a gamble and insurance. The insurance case has the same parameters as the high payment case and the gamble has the same parameters as the low payment case. Since a within-subject design is used, the case order was varied and subjects were randomly assigned to one of the four orders.

Design Innovations

The Use of Cash Incentives in This Behavioral Study

To insure that the response of subjects are economic choices, this study makes use of cash incentives. In behavioral experiments which lack cash incentives, the subjects all experience the same outcome regardless of their

choices. By contrast, in this design, a subject who selects a riskier option will get a different outcome from that of a subject who selects a less risky option.

The use of cash incentives gives new vitality to the question of external validity in experimental design. The question is whether or not the results of an experiment can be generalized to the behavior the experiment was created to describe. Specifically, the economic decisions made in the experimental task may not generalize to real tax filing judgments by taxpayers. If subjects respond to the cash incentives, the responses collected have external validity only if the experimental incentives are meaningfully related behavior of interest. Therefore, the experimental incentives tend to become the focus of all concerns about external validity of the experiment.

Berg and Dickhaut (1986) and Davis and Swensen (1988) rationalize the extension of cash incentives in lab market research to behavioral experiments in accounting and taxation. They emphasize direct tests of economic models (Berg and Dickhaut, 1986, p. 10; and Davis and Swensen, 1988, p. 42). Their emphasis on tests of theory is a response to the external validity issue. The incentives in economic models such as the agency model are generally simple enough so that they can be convincingly implemented experimentally. The experimental task thus imitates the phenomenon under study because a model is being studied. The experimental

results, it is argued, can be generalized only through the general applicability of the theory (Davis and Swensen, 1988, p. 48). This line of thought also leads to the conclusion that the experimental tasks should be abstracted from context (Mead, 1988) and that persons inexperienced in real-world instances of the judgments under study might be preferred subjects (Davis and Swensen, p. 51). Applications other than tests of theories have not been as fully rationalized (Davis and Swensen, 1988, pp. 44-45; and Plott, 1982, pp. 1520-1521).

This study addresses a question which is specific to a context. The question is how theories of decision under risk, expected utility theory and prospect theory, apply to tax filing choices. The theories are accepted as maintained hypotheses and are not tested. The experimental tasks have not been stripped of contextual references. On the contrary, context is introduced as a manipulated variable. Finally, subjects experienced in the behavior under study were considered essential to the design. Clearly, a new approach to the use of cash incentives in tax research is being advocated.

The approach advocated is pragmatic and is based on several assumptions. Experimentation is here assumed to be a valid research method, so the fact that external validity is an endemic problem is accepted as a limitation of experimentation, not a prohibition against experimentation.

It is also assumed that the use of cash incentives neither creates nor eliminates the general problem of external validity. From this perspective, the experimental incentive structure does not have to bear the burden of perfectly imitating life. Instead, cash incentives may diminish or increase the external validity of the experimental results. The final assumption is that cash incentives should be evaluated by the effect of the incentives in making the problem of external validity more or less severe.

The effects of the cash incentives on external validity can then be defined. Cash incentives increase the severity of the external validity problem if the cash incentives create responses which are less informative about the behavior under study than the responses obtainable in the absence of such incentives. Cash incentives diminish the severity of the external validity problem if the cash incentives create responses which are more informative about the behavior under study than the responses obtainable in the absence of such incentives. No matter how informativeness is operationalized, neither condition can be objectively proved, but note that neither can it be proved that an experiment has perfectly implemented the economic incentives assumed in an economic model (Davis and Swensen, 1988, p. 43). Therefore nothing is lost through this admission that the cash incentives do not imitate life (even if life is only an economic theory). On the other hand, we gain a focus which

associates the incentives with the traditional issues in behavioral experimentation of obtaining responses relevant to describing the behavior of interest.

The reason for bringing the economic variables into the task is to measure the subject's behavior, not just his attitude. Leaving to others the debate about the distinction, if any, between attitude and behavior, the decision context and decision task are both ones for which taxpayer self-insight is highly questionable. The taxpayer's attitude toward a hypothetical filing can be no more accurate than his ability to report his expected action. In the first place, people have notoriously limited self-insight into their own risk preferences. In the second place, there is evidence that taxpayers cannot accurately report their own past compliance behavior. If taxpayers cannot accurately report what they have already done, then certainly their hypothetical predictions of what they would do must be viewed with caution. By bringing cash incentives into the experimental task, the responses are behaviors, and they may well provide more accurate clues to the taxpayer choice under uncertainty than the taxpayer's expressed attitude to hypothetical situations.

The cash incentives arrangement here ease the cognitive burden of performing the experimental task. The task is to choose a preferred lottery from among five complex lotteries. The lotteries in the cash payoffs are played out with a

spinner, and experimental cash flows reflect the costs and benefits associated with the various options. A spinner with wedges proportioned to implement risk levels allows a simple visualization of variations in risk. The experimental payments and refunds allow the subject to see his cash balance by looking at the size of his fund. Another likely benefit from the use of cash incentives is to involve the subjects in comparing complex lotteries. Comparing lotteries is a task requiring considerable effort. The cash incentives may result in choices which are the product of greater cognitive attention and effort than would have occurred otherwise. The cash incentives in this task, therefore, should result in responses more characteristic of the attention, concentration, and evaluation of uncertain options characteristic of tax filings than would be the case without the cash incentives.

The cash incentives also constitute an important control. They implement the relevant parameter values for decisions under risk identically across cases and contexts. If context must be controlled, an alternative to banishing context is to include a variety of contexts. Varying contexts, however, may produce subjective variation in the perceived risks. If the cash payoffs across contexts are objectively identical lotteries, then variation in choice from one context to another cannot be the result of subjective estimations of context-specific risks. Also, the

cash incentives are aligned with tax liability minimization. The cash incentives, therefore, implement the requirement stated earlier for a PR effect, namely that it significantly affect the amount of liability reported.

Finally, the cash incentives provide a strong motive for subjects to choose options based on their own preferences. If subjects have the opportunity to please the experimenter, they sometimes will do so, a result sometimes called a demand effect. A repeated-measures design discloses the treatment differences to the subjects, making a repeated-measures design particularly prone to the appearance of demand effects on the responses of subjects. The cash incentives within the experiment provide a strong motive for subjects to focus on their own preferences instead of trying to imagine a response the experimenter might wish to receive. This means that a within-subject design with cash incentives is less likely to be tainted by experimentally induced demands than would otherwise be the case.

The Within-Subject Design

The cash incentives introduce a strong motive for subjects to ignore induced demands across multiple treatments, making a within-subject design feasible. The theory is that choices are different when risk preferences change. The mechanism which, according to theory, prompts the risk preference shifts by individuals is facing a lottery with positive outcomes as compared to a lottery with negative

outcomes. The level of risk an individual is willing to accept changes within the individual. Therefore, the judgment under study is one in which a within-subject approach more closely matches the mechanism at work than does a between-subject approach (Keren and Raaijmakers, 1988).

A within-subject design also allows a much better chance of detecting the effect than does a between-subject design. The increased statistical power achieved in most within-subject designs is enhanced here because risk preferences are variable from one individual to the next. The within-subject design allows for statistically filtering out that variability so as to increase the power of the statistical tests.

The Context Manipulation

A context manipulation allows a different approach to the control of contextual effects than does abstracting the experiment from all context. The way to show that results do not depend on context is to show that the results do not vary from one context to another. Instead of banishing contexts, multiple contexts should be introduced. If context does matter, then the context manipulation will supply insight into how tax filing judgments differ from other risky choices, which is precisely what we want to know.

The context manipulation also allows for a check on the external validity of the experiment. A common criticism of the use of cash incentives in experimental settings is that

subjects may respond to the experimental incentives differently from the way the same subjects would respond to real-world choices. If that occurs, the use of cash incentives within the experiment will yield responses which are actually garbled by the incentives. The emphasis on perfectly replicating theory in experimental tasks and on testing theory are, of course, a defense against this problem. The use of contrasting contexts, however, is an alternative strategy. That is, if taxpayers behave in the experimental setting in a gaming manner not typical of true tax filings, then the gaming tactics will apply to all the other contexts in the experiment. If the responses to the experimental tax cases are not real, then the responses to the health insurance and the gamble cases are not real in the same way. Context differences imply that the experimental incentives alone cannot account for the differences in the responses from one case to the next. Thus any differences in the responses between the contexts must reflect some extra-experimental influences. Context differences can also define taxpayer judgments in terms of comparative willingness to be exposed to risk.

Manipulation Checks And Controls

Uncertainty

Uncertainty is built into the tax cases to make the tax filing choices constitute decisions under risk. The risk is presented as uncertainty about one item where the uncertainty

is not resolvable at the time of filing. To find out whether subjects were accepting of the idea, a debriefing question asked the subject about the level of certainty typically achieved in tax filings. On a seven point scale, subjects respond to the statement, "Taxpayers frequently do not know the correct amount to report for some items on their income tax returns." A mean of agreement with the statement will indicate the overall plausibility of uncertainty in the cases.

In addition, the response to this question is used as an independent variable in a repeated-measures model. This yields several statistics to evaluate the effects of the subjects' beliefs about uncertainty in filing situations on their responses in the experimental setting. One statistic is the F statistic which tests for a main effect of the subjects' level of agreement with the statment on the subjects' choices in the cases. Another is the F statistic testing for the within-subject effect. The within-subject effect should still be important after taking into account the subjects' level of agreement with the statement. A final statistic of importance is the F statistic testing for an interaction between the within-subject effect and the main effect of the level of agreement with the statement. An insignificant interaction effect, insignificant main effect of the response to the question, and a significant within-

subject effect is compatible with a workable implementation of uncertainty in the experiment.

Group Effect

Five groups of subjects were tapped. Since there is considerable variety in the technique of recruiting subjects, a test for group effects is planned. The test parallels the test of the implementation of uncertainty. One statistic is the F statistic for a main effect of group on the choices made in the cases. Another is the F statistic testing for the within-subject effect. The within-subject effect should still be important after taking any group effect into account. A final statistic of importance is the F statistic testing for an interaction between the within-subject effect and the main effect of group. An insignificant interaction effect, insignificant main effect of group, and a significant within-subject effect implies that responses in the experiment are not different from one group to the others. The same analysis is performed for all five cases together and for the three tax cases only.

Experience of Subjects: Refund Rate and Years of Filings

The experience of the subjects as taxpayers and the tax filing context are key aspects of the design. One variable checked to evaluate the effect of prior experience of the subjects was their base rate for refund filings using their most recent ten years of tax filings.

In addition to the base rate for refunds, the number of years that subjects had filed tax returns was collected to assess the impact of experience on the responses to the experimental tasks. The role of experience is assessed in the same way that group and the response to the uncertainty question are, by a repeated-measures model in which the years of filing experience is entered as an independent variable. Since only experienced taxpayers were subjects, the comparisons are among levels of experience. The groupings are taxpayers with 4 to 10 years of filings, those with 11 to 20 years of filings and those with more than 20 years of filings. The key statistics are the F statistic testing for a significant effect of experience, the F test for the within-subject effect, and the F test for an interaction between experience and the within-subject effect.

Case Order

The within-subject design has potential flaws stemming from the multiple observations from a single subject. A practice effect, the improvement in responding to subsequent cases from practice on antecedent cases, can be, and has been, controlled by varying the order of cases. Subjects were randomly assigned to one of four case orders. Violations of the homogeneity of variance assumption have more serious impact on the integrity of the statistical tests in within-subject designs than in between-subject designs. Standard statistics (Greenhouse-Geisser epsilon and Huynh-Feldt

epsilon) estimating the effects of violations of the homogeneity of variance assumption will be used to adjust probabilities associated with each F statistic in a fashion least favorable to rejection of the null hypothesis. In a paired comparison, however, the homogeneity assumption is moot (Keppel, 1982, pp. 368-372).

A differential carry-over effect cannot be controlled. Suppose that the sequence of judgments is a learning process and that the cases vary in difficulty. It is plausible that subjects who receive the easy cases first will learn differently than will subjects who receive the hard cases first. That is, the learning may depend on the order of the cases.

In this experiment, the overall task partakes of an allocation of a fund. The sequence effect here might be thought of as the effect of antecedent spending choices on subsequent spending choices. Alternatively, the overall task might be thought of as a portfolio selection task in which antecedent choices of risk levels affect subsequent ones. Differential learning might mean, for example, that receiving the gamble before other cases affects all subsequent choices differently than receiving any other case first.

The reason order effects are of such concern is that repeated-measures designs require that an order for the cases exists. The within-subject comparisons depend on the changes in a subject's responses from one case to the next. If it

turns out that the order of the cases drives the changes in responses, then the within-subject effect may be confounded with an order effect (Keppel, 1982 p. 377).

In this experiment, differential learning would appear as an interaction between the sequence of the cases the subject received and the within-subject effect. In a repeated-measures model in which case sequence is an independent variable, the F statistic is a test for a significant interaction between case order and the within-subject effect. Any significant interactions then require additional analysis.

Implications Of Results For Utility Representations

Introduction

The purpose of this study is to test hypotheses regarding the PR effect on taxpayer filing judgments. Retrieving the subjects' preferences is sufficient to test these hypotheses. Even though the construction of utility representations from the preferences was not essential to the tests of the hypotheses, the implications of the empirical results for representations of preference orderings using utility functions is important. For one thing, utility representations are well known, and therefore describing the results in terms of such representations is a good way to communicate the results. Also, most empirical studies of utility or value representations use purely hypothetical

lotteries, while this study retrieved preferences over lotteries with cash payoffs.

Risk Preference As A Dichotomy

The preference for risk reflected in the tax choices is described in terms of a dichotomy of aversion versus attraction to risk. Fixing the responses as risk-averse versus risk-loving is done first by measuring the absolute risk levels in the preferred options. In each case, the subject selected a preferred lottery from among five options. One lottery had a certain outcome, and the other four were increasingly risky, mean-preserving spreads of the certain option. To the extent that subjects bunch at either end of the scale, a general preference for risk-averse or risk-seeking choices may be gauged.

Another approach to assessing risk preferences in the tax cases is to evaluate the shapes of the utility representations implicit in the responses to the tax cases. In terms of utility representations, risk-averse behavior is associated with a utility function with a slope of less than one and a shape such that the slope declines in the parameter (e.g. the ending wealth or amount at issue), which means a concave function. Risk-loving behavior is associated with a function with a slope greater than one and a shape such that the slope increases in the parameter, which means a convex function. While the absolute slope is not measurable here, the shape of a function across the three tax cases is

available. Thus a finding of concave shapes as compared to convex shapes is compatible with risk-averse preferences. It is expected that the tax filing choices will be better described to be risk-averse than risk-loving.

Empirical Shapes Compared with Theoretical Shapes

The shape of the utility representation in the three tax cases can also be evaluated to see whether a shape considered typical in expected utility theory or the s-shaped value function hypothesized in prospect theory better fits the empirical responses. A statistically significant PR effect does conform with the s-shaped value function of prospect theory, in which the function is steeper in losses than in gains and in which the inflection point is near zero. This result, however, is not unambiguously in conflict with the concave utility function with first derivative positive and second derivative negative, which is a risk-averse function compatible with expected utility theory.

Figure 4-2 is a representation of the refund, the low payment, and the high payment lotteries fixed on the typical s-shaped value function hypothesized by Kahneman and Tversky. This representation assumes that the settlement amount is the frame of reference used in the three tax cases. The cases differed only in the prepayment level, but the experimental task was only to pay the settlement amount (and any costs from understatement). The cash value of the cases in payoff dollars differed since only the settlement had to be made.

The refund case resulted in a net addition to the starting fund, while the high payment case required a diminution of the fund.

Since subjects chose from among five lotteries in each case, each of the three cases is represented as a range, rather than a point. Figure 4-3 is a representation of the high payment, the low payment, and the refund cases fixed on a concave, risk-averse utility function which is hypothesized to represent the most typical ordering of preferences according to expected utility theory. The feature distinguishing the value function and utility function representations is the frame of reference. The frame of reference for valuing the cases is the amount at issue in the case of the value function and the ending wealth position in the case of the utility function. From this difference, it follows that the utility function is always in a positive quadrant, while the value function has both a positive and negative quadrant. Kahneman and Tversky hypothesize a steeper slope for the value function for a given negative amount than for the same amount in the positive quadrant. The utility function is pictured typically here, and it is also steeper in the negative amount (high payment) than in the refund. Therefore, a difference in the risk level preferred in the high payment and refund cases is compatible with both the value function and the utility function.

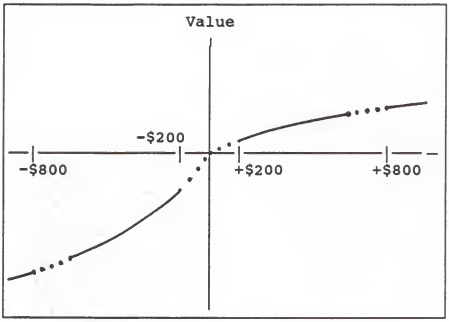

Figure 4-2
Hypothetical Value Function
With Tax Cases Shown As Dotted Portions of Function

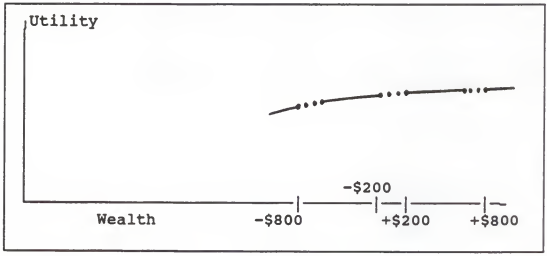

Figure 4-3
Hypothetical Concave Utility Function With Tax Case
Amounts Shown As Dotted Portions of the Function
(Ending Wealth = Beginning Wealth Plus Amount At Issue)

With only three empirically observed levels for P, assessing shapes of functions is tenuous. However, the frequencies for the four possible shapes (linear, convex, concave, and s-shaped) can be compared to frequencies to be expected by chance to see whether or not a modal pattern is exemplified in the empirical data. A convex shape means that a riskier choice was made for either the low payment or the refund than in the high payment case, and the choice in the refund case was at least as risky as the choice in the low payment case. A linear shape means that the same option was selected in all three tax cases. A concave shape means that a less risky choice was made in either the low payment or refund case than in the high payment case and that the choice in the low payment case was at least as risky as the choice in the refund case. If the shape is not linear, convex or concave, then it is s-shaped. An s-shape means that the choice of the risk level preferred changes in one direction between the high payment and the low payment cases and changes in the opposite direction between the low payment and the refund cases.

The shape for the aggregated responses can be assessed through orthogonal polynomial contrasts. With only three tax cases, the quadratic is the highest order polynomial which can be tested. The cubic order is necessary to test for an s-shape, so the polynomial can only distinguish significant curvature from linearity. The test for this is

an F test for the within-subject quadratic contrast. A significant F statistic for the quadratic contrast implies curvature. The shape can also be evaluated by the frequencies with which subjects made choices which implied each of the four shapes. The only expectation is that, in parallel with the expectations of generally risk-averse choices, a convex shape will be observed less frequently than would be expected by chance. The chance frequencies assume random choices in the three cases. A binomial test using the appropriate probabilities can be used to test the significance levels of observed frequencies for the different shapes.

The utility and value representations might yield conflicting predictions for the choices in the low payment case. In a typical utility function ($U'>0$ and $U''< 0$), like the one pictured in Figure 4-3, the slope of the function for the low payment case should lie in a range consisting of the slopes for the low payment and refund cases. The same is not necessarily true of the value function. The low payment case is neither unambiguously positive or negative. The value function has an inflection point between the positive and negative quadrants making the slope of function around the inflection point unpredictable. If the settlement amount constitutes a frame of reference, then the low payment case might prompt choices involving risks greater than the risk preferred in either the refund or low payment case. Of course, given the reflex point in the region, the possibility

that the risk level preferred in the low payment case might fall within the range of the risk levels preferred in the high payment and refund cases could not be rejected. The rates at which the risk level preferred in the low payment case falls within, and does not fall within, the range of risk levels in the refund and high payment case can be compared to rates expected by chance. A binomial test using the appropriate probabilities can be used to test the significance levels of observed frequencies. No expectation exists for the results.

Context Effects and Utility Representations

The context effects hypothesized here are incompatible with both utility and value function representations. The lotteries in the low payment and gamble cases are identical as are the lotteries in the high payment and health insurance cases. If the null is rejected in Hypothesis 3, then there are two observations at a single point on the horizontal axis which will have significantly different slopes on a function. A functional representation therefore fails. It also means that the preferences are intransitive because, for instance, in two identical lotteries, B is preferred to C in one, but C is preferred to B in the other. This is the same pattern of inconsistency prompted by framing identical lotteries as either gains or losses. Except for the general emphasis on editing phases in which frames of reference influence risk preferences, preference reversals across contexts are not

predicted in prospect theory. However, the notion of context-specific risk preferences is more in the spirit of prospect theory than in the spirit of expected utility. The tests for Hypothesis 3 described earlier constitute tests for significant within-person differences in risk preferences due to context.

CHAPTER 5

RESEARCH RESULTS

Introduction

The experimental manipulations produced statistically significant results. The settlement level, P, had a significant effect on (1) the amount of an uncertain item reported in the tax filing cases and (2) on the risk level selected in the tax filing cases. The context, CON, had a significant effect on the level of risk preferred by subjects, and preference reversal occurred between the tax and nontax contexts. The risk level preferred in the gamble was significantly greater than that preferred in all tax cases, and the risk level preferred in the health insurance case was significantly lower than in the high payment tax case. The manipulation and control checks disclosed no severe design problems. The utility and value representations implicit in the responses to the tax cases were risk-averse, as opposed to risk-seeking, and were not incompatible with either a concave utility function or an s-shaped value function. The context manipulations produced results incompatible with functional representations.

The Hypotheses

Hypothesis 1

The first hypothesis involves the effect of the settlement amount on the amount of an uncertain item reported in the tax filing task.

Ho_1: There is no significant PR effect on the amounts reported in the tax filings.

Hypothesis 1 is rejected. The hypothesis of no difference is rejected with an alpha value of less than one percent. An analysis of variance appears in Table 5-1, and the analysis for all three tax cases together appears in Panel A. COM was not significantly different in the refund and the low payment cases, and COM in the high payment case differed from COM in both the refund and the low payment cases. Panel B of Table 5-1 gives the analysis of variance for the comparison of COM in the refund and the high payment cases. Panel C of Table 5-1 is the analysis of variance for the comparison of COM in the low payment and the high payment cases. In all these comparisons, the hypothesis of no effect can be rejected at an alpha level of less than one percent.

The direction of the differences is as expected. That is COM, the amount included in taxable income, was lower in the high payment case than in the low payment and the refund cases. The means for COM, as well as the means for other variables, appear in Table 5-2. The mean responses for COM were 69%, 70%, and 57% in the refund, the low payment and the high payment cases respectively.

Table 5-1
Within-Subject Effect of P (Settlement Amount)
On COM (Percentage of Item in Question Reported)
For Tax Cases

Panel A: All Tax Cases

Source	DF	Type III SS	F	PR>F
Within-Subject				
Case	2	.658	5.84	.004
Error	128	7.217		

Panel B: Refund and High Payment

Source	DF	Type III SS	F	PR>F
Within-Subject				
Case	1	.462	8.28	.0054
Error	64	3.569		

Panel C: Low Payment and High Payment

Source	DF	Type III SS	F	PR>F
Within-Subject				
Case	1	.524	8.23	.0056
Error	64	4.070		

Table 5-2
Mean Responses to Variables

NAME	DESCRIPTION	N	MEAN
	PERCENT OF MAXIMUM PAYMENT PAID (0%-100%)		
COM1	HEALTH INSURANCE	65	0.685
COM2	TAX REFUND	65	0.688
COM3	TAX LOW PAYMENT	65	0.696
COM4	TAX HIGH PAYMENT	65	0.569
COM5	GAMBLE	65	0.385
	RISK LEVEL SELECTED (0%-40%)		
R1	HEALTH INSURANCE	65	0.146
R2	TAX REFUND	65	0.155
R3	TAX LOW PAYMENT	65	0.151
R4	TAX HIGH PAYMENT	65	0.202
R5	GAMBLE	65	0.268
	RISK LEVEL SELECTED, RANK ORDER (0%-40%)		
RK1	HEALTH INSURANCE	65	0.187
RK2	TAX REFUND	65	0.193
RK3	TAX LOW PAYMENT	65	0.204
RK4	TAX HIGH PAYMENT	65	0.249
RK5	GAMBLE	65	0.297
Q1	TAXPAYERS ARE OFTEN UNCERTAIN, (1=AGREE AND 7=DISAGREE)	65	2.754
Q5	REFUND FILINGS PERCENTAGE	65	0.843
Q33	NUMBER OF TAX FILINGS, (2=4-10, 3=11-20, AND 4=MORE THAN 20)	65	3.231

Hypothesis 2

Hypothesis 2 also involves the PR effect. This hypothesis has to do with the effect of the settlement amount on the level of risk preferred for the item at issue.

Ho_2: There is no significant PR effect on the level of risk preferred for reporting an item at issue.

The level of risk is measured by the variable R. An analysis of variance is shown in Table 5-3 for the effect of P on R. Panel A gives the analysis of variance for all three tax cases together, Panel B gives the comparison of the effect in the refund and the high payment cases only, and Panel C gives the analysis of variance in the low payment and the high payment cases only. The hypothesis of no effect is rejected for all three comparisons at an alpha level of one percent. As was true in the test of Hypothesis 1, the differences are in the hypothesized direction. The mean value for R was 16%, 15%, and 20% for the refund, the low payment, and the high payment cases respectively. The means for R are given in Table 5-2.

Hypothesis 3

Hypothesis 3 involves the variation in contexts. There are two versions of this hypothesis, one for each of the nontax contexts. Hypothesis 3A involves the contrast between the health insurance and tax contexts.

Ho_{3A}: There is no significant difference in the level of risk preferred in a tax filing and in a purchase of health insurance.

Table 5-3
Within-Subject Effect of P (Settlement Amount)
On R (Risk Preferred For Item in Question)
For Tax Cases

Panel A: All Tax Cases

Source	DF	Type III SS	F	PR>F
Within-Subject				
Case	2	.106	5.00	.0082
Error	128	1.360		

Panel B: Refund and High Payment

Source	DF	Type III SS	F	PR>F
Within-Subject				
Case	1	.073	7.05	.0100
Error	64	.662		

Panel C: Low Payment and High Payment

Source	DF	Type III SS	F	PR>F
Within-Subject				
Case	1	.086	7.74	.0071
Error	64	.710		

Hypothesis 3A is rejected. The hypothesis of no difference is rejected at an alpha value of less than one percent. An analysis of variance for this comparison appears in Panel A of Table 5-4. The direction of the differences is as expected. That is R, the risk level preferred, was lower in the health insurance case than in the high payment case. The means for R were 15% and 20% in the health insurance and the high payment cases respectively (see Table 5-2). Since the options in the health insurance and the high payment cases were identical sets of lotteries, the rejection of Hypothesis 3A implies preference reversal. That is, from among identical options, a riskier option was preferred in the tax context and a less risky option was preferred in the health insurance context.

There was no statistically significant difference in the risk level preferred in the health insurance case and in either the refund or the low payment cases. Panel B of Table 5-4 is the analysis of variance testing a significant difference in R for these three cases (health insurance, refund, and low payment).

The other part of Hypothesis 3 involves the contrast between the tax and the gamble contexts.

Ho_{3B}: There is no significant difference in the level of risk preferred in a tax filing and in a gamble.

Hypothesis 3B is rejected. An analysis of variance for this comparison appears in Panel A of Table 5-5. The direction of

Table 5-4
Within-Subject Effect of P (Percent of Maximum Payment Paid) on R (Risk Preferred For Item in Question) For Health Insurance Versus Tax Cases

Panel A: Health Insurance versus High Payment

Source	DF	Type III SS	F	PR>F
Within-Subject				
Case	1	.104	7.25	.009
Error	64	.919		

Panel B: Health Insurance, Refund, and Low Payment

Source	DF	Type III SS	F	PR>F
Within-Subject				
Case	2	.003	.09	.909
Error	128	1.883		

Table 5-5
Within-Subject Effect of P (Percent of Maximum Payment Paid) on R (Risk Preferred For Item in Question) For Gamble Versus Tax Cases

Panel A: Gamble Versus Low Payment

Source	DF	Type III SS	F	PR>F
Within-Subject				
Case	1	.448	36.97	.0001
Error	64	.775		

Panel B: Gamble versus High Payment

Source	DF	Type III SS	F	PR>F
Within-Subject				
Case	1	.142	10.64	.0018
Error	64	.851		

Panel C: Gamble versus Refund

Source	DF	Type III SS	F	PR>F
Within-Subject				
Case	1	.418	27.25	.0001
Error	64	.981		

the differences is as expected. That is R, the risk level preferred, was lower in the low payment tax case than in the gamble. The means for R were 15% and 27% in the low payment and the gamble cases respectively (see Table 5-2). Since the lotteries in the low payment and the gamble cases were identical, the rejection of Hypothesis 3B implies preference reversal.

The level of risk selected in the gamble was significantly greater than in all the other cases, including the high payment tax case. Panel B of Table 5-5 is the analysis of variance testing for a significant difference in R for the high payment tax case and the gamble. The mean for R in the high payment case was 20%, and the mean for R in the gamble was 27% (see Table 5-2).

Manipulation Checks And Controls

Uncertainty

Subjects found the uncertainty about the correct amount of an item to report plausible. On a seven point scale in which a one meant strongly agree and a seven meant strongly disagree, the average response was 2.8 to the statement, "Taxpayers frequently do not know the correct amount to report for some items on their income tax returns." Since a four was labelled as neither agree nor disagree, the mean response indicated some moderate level of agreement with the statement.

The level of agreement with this statement did not affect the responses to the cases. When the response to this statement was used as an independent variable in a repeated-measures model, (1) the main effect was insignificant, (2) the interaction between the within-subject effect and the main effect was insignificant, and (3) the within-subject effect remained significant. An analysis of variance for these tests appears in Table 5-6.

Group Effect

The test for group effects disclosed no significant main effect of group nor any significant interaction between group and the within-subject effect. The same analysis was performed for all five cases together and for the three tax cases together. An analysis of variance is in Table 5-7. Panel A is the analysis of variance for all five cases, and Panel B is the analysis for the three tax cases.

Experience: Refund Rates And Years of Filings

One variable checked to evaluate the effect of prior experience of the subjects was the proportion of the previous tax filings, up to ten years back, which were filings for refunds. The mean response was 84%, so the base rate for refunds in the subject pool is very similar to the aggregate rates described in the Background section.

The number of years of tax filings by the subjects was also collected to assess the impact of experience in the responses to the experimental tasks. The analysis of

Table 5-6
Test for the Effect of Q1 (Level of Agreement With Statement That Taxpayers Frequently Do Not Know The Correct Amount to Report) On COM (Percent of Maximum Payment Paid)

All Tax Cases.

Source	DF	Type III SS	F	PR>F
Between-Subject				
Q1	5	.098	.12	.988
Error	59	9.939		
Within-Subject				
Case	2	.720	6.28	.0028
Case*Q1	10	.451	.79	.641
Error	118	6.765		

Table 5-7
Test for the Effect of Group
On R (Level of Risk Preferred)

Panel A: Group and All Cases

Source	DF	Type III SS	F	PR>F
Between-Subject				
Group	4	.228	1.76	.148
Error	60	1.937		
Within-Subject				
Case	4	.705	12.57	.0001
Case*Group	16	.148	.66	.832
Error	118	3.364		

Panel B: Group and All Tax Cases

Source	DF	Type III SS	F	PR>F
Between-Subject				
Group	4	.178	1.70	.162
Error	60	1.570		
Within-Subject				
Case	2	.134	6.23	.0027
Case*Group	8	.066	.77	.630
Error	120	1.293		

variance for evaluating the effect of experience appears in Table 5-8. Panel A shows the effect for all cases, while Panel B shows the results for the tax cases only. The main effect of experience is significant at an alpha level of less than two percent for both the tax cases and for all cases together. The within-subject effect, however, remained highly significant, and the interaction between the within-subject and main effect was not statistically significant.

The statistics in the analysis of variance indicate that the absolute responses selected vary with taxpayer experience, but the change from one case to another occurs among both the more and the less experienced taxpayers. Table 5-9 displays the mean response of R by filing experience. In each experience group, the risk level preferred increases from the refund case to the high payment case. In each tax case, the risk level preferred increases with experience. Even though more experienced taxpayers selected riskier filing options than less experienced taxpayers, taxpayers in all experience levels tended to choose riskier options in the high payment case than in the refund case.

Case Order

The between-subject effect of case order was not statistically significant. This was expected due to the random assignment of subjects to the four different case orders. Panel A of Table 5-10 contains the analysis of variance for the effect of case order in the three tax cases.

Table 5-8
Test for the Effect of Q33 (Filing Experience
On R (Level of Risk Preferred)

Panel A: Filing Experience and All Cases

Source	DF	Type III SS	F	PR>F
Between-Subject				
Experience	2	.286	4.71	.0124
Error	62	1.879		
Within-Subject				
Case	4	.755	13.65	.0001
Case*Experience	8	.086	.78	.622
Error	248	3.426		

Panel B: Filing Experience and Tax Cases

Source	DF	Type III SS	F	PR>F
Between-Subject				
Experience	2	.259	5.39	.0069
Error	62	1.489		
Within-Subject				
Case	2	.112	5.20	.0068
Case*Experience	4	.021	.49	.747
Error	124	1.339		

Table 5-9
Means For R In Tax Cases By Years of Tax Filing Experience

YEARS OF FILINGS	N	REFUND	LOW PAYMENT	HIGH PAYMENT
4-10	14	0.09	0.11	0.16
11-20	22	0.13	0.13	0.20
>20	29	0.21	0.19	0.23

Table 5-10
Test for the Effect of SEQ (Order of Cases)
On R (Level of Risk Preferred)

Panel A: Sequence and Tax Cases

Source	DF	Type III SS	F	PR>F
Between-Subject				
Sequence	3	.064	.77	.517
Error	61	1.684		
Within-Subject				
Case	2	.107	5.20	.0069
Case*Sequence	6	.100	1.62	.1472
Error	122	1.259		

Panel B: Sequence And Health Insurance Contrast

Source	DF	Type III SS	F	PR>F
Between-Subject				
Sequence	3	.044	.59	.6260
Error	61	1.533		
Within-Subject				
Case	1	.101	7.44	.0083
Case*Sequence	1	.092	2.26	.0902
Error	61	.827		

Panel C: Sequence And Gamble Contrast

Source	DF	Type III SS	F	PR>F
Between-Subject				
Sequence	3	.027	.39	.763
Error	61	1.399		
Within-Subject				
Case	1	.458	44.54	.0001
Case*Sequence	3	.149	4.82	.0045
Error	61	.627		

There was no significant differential carry-over effect in the three the tax cases, so the order with which the tax cases were decided does not affect the refund and the high payment comparisons. There is, however, a significant interaction between the case sequence received and the within-subject effect for the context comparisons. Panel B of Table 5-10 is the analysis of variance for the comparison of the health insurance and tax contexts. Panel C of Table 5-10 is the analysis of variance for the comparison of the gamble and the tax contexts. The interaction effect for the comparison of the low payment tax case and the gamble is significant at less than one percent.

There is a very slight trend toward riskier choices in later cases. Table 5-11 gives the mean responses to the cases by position. A regression of the means for R (risk level of preferred option) against position number (where the first position is a 1, the second position 2, and so forth), yields upward slopes for all four cases (Table 5-12). The slopes, all having values of one or two percent, are not statistically significant.

Figure 5-1 is a graph of R for the gamble and the low payment cases by case position. Each point is the mean R of the 16 subjects who received the case in that particular position. The volatility in the graph reflects the substantial variation that is typical of risk preferences, making it difficult to interpret conclusively. The slight

Table 5-11
Means for R By Position of Case in Sequence

Position	Low Payment	High Payment	Health Insurance	Gamble
First	.103	.160	.125	.217
Second	.147	.236	.099	.295
Third	.184	.202	.191	.323
Fourth	.169	.211	.171	.242

Table 5-12
Simple Regression Of Case Against Position

	Low Payment	High Payment	Health Insurance	Gamble
R square	0.7397	0.2360	0.4992	0.0755
St. Error	0.0099	0.0151	0.0163	0.0255
DF	2	2	2	2
Beta	0.0235	0.0119	0.023	0.0103
T Stat.	2.38	0.79	1.41	0.40

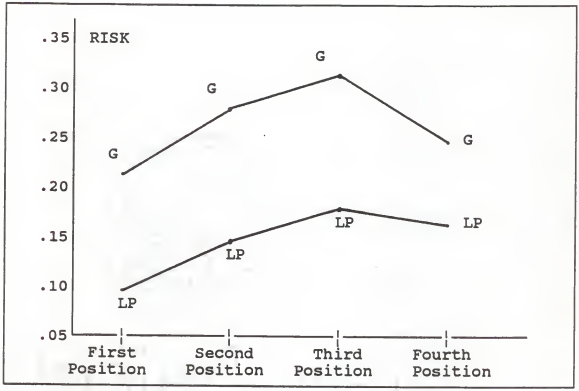

Figure 5-1
Low Payment Versus Gamble Comparison
Interaction Of Risk (R) With Case Position

upward trend is apparent in both variables, but nonparallel or crossing lines, which indicate interactions, are not apparent. Though position makes a difference, the choices differ in the gamble and the low payment cases in every position. In addition, there is no significant interaction between the sequence of the cases and the within-subject effect for the gamble compared to the high payment case.

The interaction between position and the within-subject effect for the health insurance versus the high payment tax case is not terribly strong. The interaction is significant at an alpha level of .09. Figure 5-2 is an graph of the mean responses to R by position for these two cases. Again the upward drift and volatility are apparent. Whether or not there is convergence is arguable, but the lines do not cross.

Subjects also prepared a ranking of the cases by the relative risk level preferred in the cases. Data from this ranking suggest that case order does not create the within-subject differences. After completing the five cases, subjects were asked to look back at all five cases and rank them by using each of the five risk levels once and only once across the five cases. This was a simultaneous ranking, and was therefore, by construction, not subject to any order effect. The ratings of the health insurance and high payment case from this ranking were evaluated to see whether the within-subject effect would remain significant when the judgment was a simultaneous one. Using the health insurance

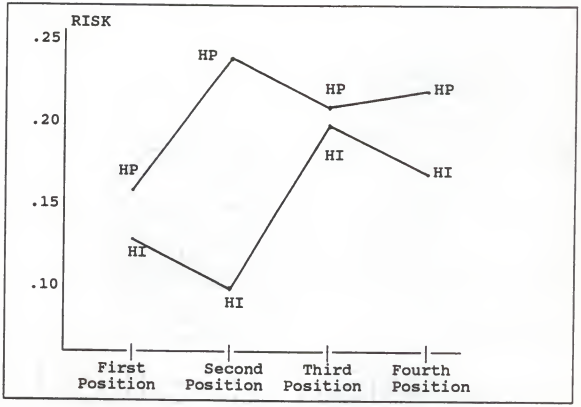

Figure 5-2
High Payment Versus Health Insurance Comparison
Interaction Of Risk (R) With Case Position

and the high payment responses from this ranking in the repeated-measures model with sequence as an independent variable yielded the same highly significant within-subject effect but an insignificant interaction between the within-subject effect and the sequence with which the cases were received.

A reasonable case can be made that sequence has not produced the context effects. The graphs of the interactions are volatile, but demonstrate no statistically measurable differences in slope. Order appears to make a difference in the judgments, but it does not negate the context differences. There may be an anchor-and-adjustment process at work in the selection of specific risk amounts from case to case, and the drift toward riskier choices in later cases is satisfying as it parallels the diminution of the subjects' spending fund. The absence of significant interaction in the tax cases shows that order had no impact on the PR effect in the tax setting.

Implications Of Results For Utility Representations

Risk Preferences As A Dichotomy

The choices made by the subjects in the three tax cases are generally risk-averse. The mean choices for the refund, the low payment, and the high payment cases respectively were 16%, 15%, and 20%. The mean response was just one step up from the certain option. Only two subjects selected the riskiest option in all tax cases, so 63 out of 65 subjects

had riskier options available than were selected. The modest level of the risk, never even attaining 50%, and the preference for lower risk options across mean-preserving spreads suggests risk-averse behavior.

Empirical Shapes Compared With Theoretical Shapes

The mean choices of risk level in the refund and the low payment case were essentially the same while the risk level preferred in the high payment case was significantly higher. These mean responses in the tax cases are displayed on a stylized utility function in Figure 5-3 and on a stylized value function in Figure 5-4. These figures are drawn with a straight line through the refund and low payment regions to represent the same slope to the utility and value function representations in these two cases. The mean choice of risk level in the three tax cases is not incompatible with either functional representation.

The stylized figures represent means. The frequency with which the four possible shapes occurred in the responses to the tax cases can also be counted. Only six subjects' preferences yield a convex function where convex means that a riskier choice was made in either the low payment or the refund (or both) than in the high payment case and that the choice in the refund case as at least as risky as in the low payment case. By contrast, 20 of the subjects made choices yielding a concave function, where concave means that a less risky choice was made in either the low payment or the refund

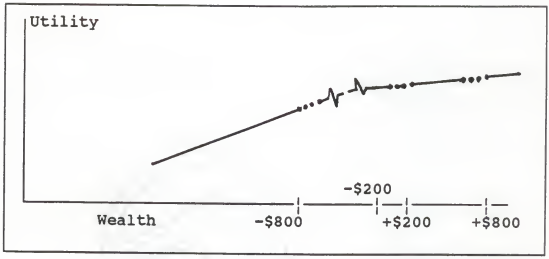

Figure 5-3
Stylized Empirical Utility Function With Tax Case
Means Shown As Dotted Portions of the Function
(Ending Wealth = Beginning Wealth +/- Amount At Issue)

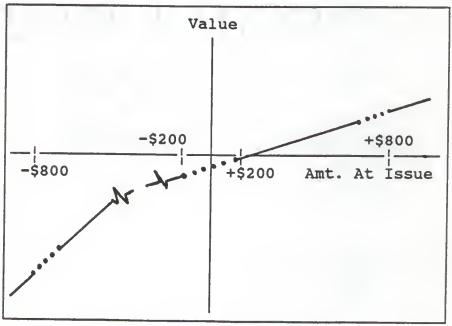

Figure 5-4
Stylized Empirical Value Function
With Tax Case Means Shown As Dotted Portions of Function

case than in the high payment case and that the choice in the low payment case was at least as risky as in the refund case. Seventeen subjects made the same choice in all three cases (a linear function).

The preferences which yield linear, convex, and concave utility representations account for 43 of the 65 subjects' choices. The remaining 22 subjects made choices in the tax cases which are not linear, or concave, or convex (as defined above). Instead, the risk level preferred changes in one direction going from the high payment case to the low payment case, and the risk level preferred changes in the opposite direction going from the low payment case to the refund case.

The chance frequency for each of these orderings (linear, convex, concave, and s-shaped) provides a benchmark for interpreting the empirical frequencies observed in the experiment. Table 5-13 shows the chance frequencies and the empirical frequencies for these shapes. The data in Table 5-13 shows that the empirical frequencies for the linear ordering and the convex ordering are far from the frequencies to be expected by chance. The linear ordering occurs more frequently than would be expected by chance, and the convex ordering occurs less frequently than would be expected by chance. The concave and s-shaped orderings are not far from frequencies that would be expected by chance. The results reflect high variability from one person to the next in risk

Table 5-13
Expected Random Occurrence Rates And Empirical Occurrence Rates For Linear, Convex, Concave and S-Shaped Preference Orderings Over Three Tax Cases

Preference Ordering	Expected Random Occurrence Rate	Empirical Occurrence Rate
Linear	.04	.26
Convex	.24	.09
Concave	.24	.31
S-Shaped	.48	.34

preferences. In fact, each of the 13 possible orderings occurred in at least one subject's choices.

The analysis of variance for the orthogonal polynomial contrasts appear in Table 5-14. The linear contrast was significant at an alpha level of .01, and the quadratic contrast yielded an F statistic with an alpha level of .085.

The shape of the functional representations based on aggregate measures (means and the polynomial contrast) and based on the frequency counts from individual subjects is not compatible with a convex utility representation. When coupled with the selection of options with lower risk, the information on the shapes of the representations yields a conclusion of risk-averse, as compared to risk-seeking, preferences in the tax context as a whole. Neither the mean responses, nor the counts of functional shapes implicit in the choices, nor a significant quadratic contrast are incompatible with either a risk-averse utility function or the s-shaped value function hypothesized in prospect theory.

One other test has the possibility of determining whether either expected utility theory or prospect theory is incompatible with the empirical results. This test involved the risk level preferred in the low payment case in comparison to the risk levels preferred in the high payment and the refund cases. The statistic of interest is the frequencies with which the risk chosen in the case with parameter values mid-way between the parameter values of the

Table 5-14
Test for Orthogonal Linear and Quadratic Contrast
On Within-Subject Effects of R In All Tax Cases

Panel A: Linear Contrast

Source	DF	Type III SS	F	PR>F
Within-Subject				
Mean	1	.073	7.05	.0100
Error	64	.662		

Panel B: Quadratic Contrast

Source	DF	Type III SS	F	PR>F
Within-Subject				
Mean	1	.033	3.05	.0855
Error	64	.697		

refund and the high payment case (i.e., the low payment case) fell within, and fell outside, the range of the risk preferred in the other two cases. Table 5-15 gives the frequency of choices in which the risk level chosen in the low payment case was outside the range of the risk levels chosen in the refund and the high payment cases. It also shows the frequency of choices in the low payment case that are not outside the range of the risk levels preferred in the refund and the high payment cases. The frequencies to be expected by chance are also given. The empirically observed frequency with which the rating for the low payment case was outside the range of risk levels of the refund and high payment case was lower (but probably not statistically significant lower) than the rate to be expected by chance. Therefore, using a utility representation generally believed to be typical, the results from the tax cases are not inconsistent with expected utility theory.

Nontax Choices And Typical Utility and Value Functions

The nontax cases resulted in choices of very different risk levels than did the tax cases even though the options in the nontax cases were the same as those in the tax cases. Risk preference reversals inconsistent with expected utility occurred in the choices made in one context as compared to another context, as was established through the rejection of the null in Hypothesis 3. Figure 5-5 repeats the stylized utility representation of the results in the tax cases shown

Table 5-15
Expected Random Occurrence Rates And Empirical Occurrence Rates For Preference for Risk in Low Payment Case Outside And Not Outside Range of Risk Preferences In High Payment And Refund Cases

Risk in Low Payment Case in Relation To Range of Risk in High Payment and Refund	Expected Random Occurrence Rate	Empirical Occurrence Rate
Risk in Low Payment Is Not Outside Range	.52	.67
Risk in Low Payment Is Outside Range		
Outside High	.24	.12
Outside Low	.24	.22
Outside Total	.48	.34

in Figure 5-3, but additional segments appear in the range of the low tax payment for the gamble and in the range of the high tax payment for the health insurance. The additional segments represent the mean responses to the risk level preferred in the gamble and in the insurance cases. While the options in the health insurance case were identical to those in the high payment tax case, the risk level preferred implies a utility function slope equal to that in the refund tax case. The dollar lottery in the gamble was identical to that in the low payment tax case, but the risk level preferred implies a utility function slope much steeper than that in any of the other four cases.

The preference reversals due to context are just as peculiar when pictured with a single value function as they are when pictured on a typical utility curve. Figure 5-6 shows the stylized value function with the empirical results from the health insurance and gamble cases displayed as line segments in the range of the high payment and the low payment cases respectively. The health insurance line segment, though appearing in the loss quadrant, has a slope equal to that in the refund case, while the gamble segment, though appearing in the range of the low payment tax case has a slope steeper than in all other cases.

There is nothing in the editing steps described in prospect theory predicting this particular context-induced preference reversal. Nevertheless, the preference reversal

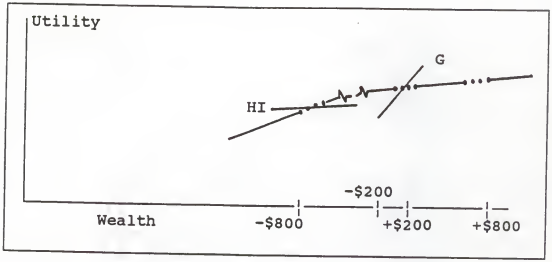

Figure 5-5
Stylized Empirical Utility Function With Tax Case
Means Shown As Dotted Portions of the Function.
Gamble and Health Insurance Means Shown
As Line Segments
(Ending Wealth = Beginning Wealth Plus Amount At Issue)

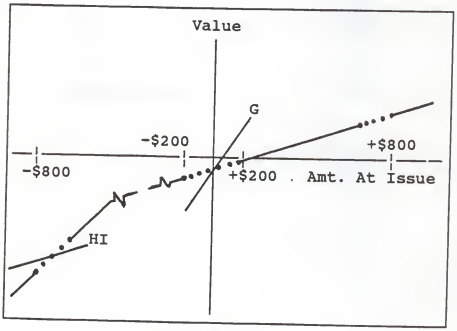

Figure 5-6
Stylized Empirical Value Function With Tax Case
Means Shown As Dotted Portions of Function
And Gamble (G) And Health Insurance (HI) Means Shown
As Line Segments

across contexts is more compatible with prospect theory than it is with expected utility theory. In expected utility theory, the objective features of the lotteries enter the valuation. In prospect theory, subjective editing alters the objective features of the prospect (e.g., the Allais Paradox). Here the value of the prospects depend on the context within which the risk and penalties occur. Even when the values of the parameters are held constant, the prospects have different values from one context to another. The implication is that functional representations are context-specific.

CHAPTER 6

SUMMARY AND CONCLUSION

Summary

The average taxpayer emerges from this study as risk-averse in all tax settings. On a comparative basis, he is almost as risk-averse in tax choices as he is in choosing health insurance, and he is much more risk-averse in tax choices than he is in choosing a gamble. In addition, this average taxpayer prefers to be overpaid at settlement. Since withholding, as opposed to estimated payments, is the prepayment system used by most taxpayers, this means that he prefers to have withholding at a level which results in overpayment. In this experiment, 49 of the 65 subjects expressed a preference for withholding levels producing overpayments. Only eight subjects expressed a preference against withholding levels producing overpayments (eight subjects were indifferent). The bias of the withholding system in the United States toward overpayment therefore conforms to the preferences of most of these taxpayers.

The subjects in this experiment have been successful in getting the result preferred. Of the 49 subjects preferring withholding producing overpayments, 41 had received refunds

in at least eight of the most recent ten years. Only three had refunds in less than half of the most recent ten years of filings. Also, the percentage of refund filings increased with the strength of the preference for being overpaid.

The bias in the withholding system makes it more difficult for wage earners to prepay less than the liability, but experienced taxpayers do evidence some control over the prepayment level. Subjects expressing a preference against a withholding level producing an overpayment did, in fact, have payment-due filings more frequently than those desiring or indifferent to overpayment. Of the eight subjects who expressed a preference against withholding levels resulting in overpayments, three had refunds in eight of the last ten years, three had refunds in six of the preceding ten years, one had three refunds in the last ten years, and one had only two refunds in the last ten years. In sum, the average taxpayer seems to be risk-averse in tax filing choices and to prefer, and to attain, a tax filing in which he is overpaid.

The settlement amount (payment or refund) had a significant effect on the amount of an uncertain item reported in the tax filing cases. The amount included in taxable income, was lower in the high payment case than in the low payment and the refund cases. Options available for the amount to report ranged from 0% to 100%. The means for

compliance were 69%, 70% and 57% in the refund, the low payment, and the high payment cases respectively.

The settlement amount (payment due or refund) also had a significant effect on the risk level selected in the tax filing judgments. Options available for the risk level ranged from 0% to 40%. The means for risk level selected were 16%, 15%, and 20% for the refund, the low payment, and the high payment cases respectively.

The context had a significant effect on the level of risk preferred by subjects, and preference reversal occurred between the tax and nontax contexts. The risk level preferred in the gamble was significantly greater than that preferred in all the tax cases. The mean risk levels were 15% and 27% in the low payment and gamble cases respectively. The risk level preferred in the health insurance case was significantly lower than in the high payment tax case. The risk level preferred in the health insurance case was the same as the level chosen in both the refund and the low payment cases. The means for the risk level selected were 15% and 20% in the health insurance and the high payment cases respectively.

The nontax cases resulted in choices of very different risk levels from those in the tax cases even though the lotteries in the nontax cases were identical to those in the tax cases. This allows a relative fix on the willingness of subjects to be exposed to risk in tax filings. The

willingness to be exposed to risk in tax filings is almost as low as in the purchase of health insurance and is much lower than in gambles.

The mean choices made by the subjects in the three tax cases are low risk choices; the average risk level for all three tax cases combined was 17%. When coupled with the selection of options with lower risk, the empirically derived shapes of the representations yields a conclusion of risk-averse, as compared to risk-seeking, preferences in the tax context as a whole.

The utility and value representations implicit in the responses to the tax cases were compatible with either a risk-averse concave utility function or an s-shaped value function. The manipulation and control checks disclosed no severe design problems.

The context manipulations also constitute a most important experimental control. The significance of the context differences means that when cash incentives within the experiment were held constant, subjects chose different options. Thus subjects are not responding solely to the experimental incentives with game-like abandon. Instead the experimental choices, despite the cash incentives, still reflect extra-experimental influences. Therefore, the key difference explored in the study, the difference between the amounts reported in refund and payment-due tax filings, very likely also reflect extra-experimental influences.

The context manipulations produced results incompatible with either an expected utility functional representation or an s-shaped value functional representation. The preference reversal across contexts is more compatible with the influence of frames of reference as described in prospect theory than it is with expected utility theory. The implication, however, is that functional representations are context-specific, so that people behave as if they had multiple preference ordering schemes.

The aggregate data reported in Chapter 1 also support this picture of the average taxpayer. The data reflect a consistent pattern of overpayment for the average taxpayer. It appears likely that part of the expertise developed with tax filing experience is the ability to control withholding so as to attain a desired settlement outcome. It strains credulity to believe that such a stable pattern is imposed contrary to taxpayer preferences. One would have to assume, for example, that taxpayers are unaware of any opportunity to alter W-4 data. If such a stable pattern was produced contrary to taxpayer preferences, one would have to credit the IRS with a remarkable feat. One would have to assume that contrary to taxpayer preferences and in the face of Congressional pressure to reduce the level of overpayments, the IRS has still managed to progressively increase the overpayment levels. It is much more plausible that most taxpayers accurately report their preference for being

overpaid at filing. This implies a preference for a decision frame associated with less risky choices which also fits perfectly with the preference for risk-averse tax filing judgments. It appears that the average taxpayer prefers both a risk-averse frame of reference and risk-averse choices.

Limitations

There are limitations on the generalization of these results. First, the representativeness of the subject pool is unknown. Absent the ability to identify a random sample of taxpayers, obtaining a random sample is impossible. The alternative chosen in this study was to diversify the subject group by recruiting subjects from diverse organizations. Generalization from the subject pool to the taxpaying population is hazardous without a random sample.

The second consideration in generalizing these results stems from the use of cash incentives. Cash incentives solve the problem of inferring economic choices from expressions of attitudes. Social and ethical beliefs influence people's utility for monetary outcomes in actual tax filing decisions and in experimental settings. Social and ethical beliefs, however, may affect a subject's utility for monetary outcomes in the lab setting differently from the way they affect a subject's utility for monetary outcomes in tax filings. The strategy chosen in this study was to use context comparisons as a control, and differences from one case to the next as

the measures of effect. The extension of the findings to actual tax filings is, however, an open question.

In addition, many parameters in the cases are held within limited bounds. For example, taxable incomes in the cases vary only from about $25,000 to $30,000, items at issue are about 15% of the liability, and the settlement amounts are not more than 15% of the total liability. Even within an experimental setting, generalization outside the range of parameter values is not possible.

Conclusion

The import of this study is that higher taxable income is reported, and higher tax liabilities are paid, when the individual taxpayer is filing for a refund than when the taxpayer is filing with a payment due. The timing of the tax payments affect the amount of the tax liability reported. This PR effect occurs because the willingness of individual taxpayers to take risky positions increases when the taxpayer has a payment due rather than a refund. The PR effect may occur whenever decision under uncertainty characterizes tax filing choices. Therefore, the PR effect should be moderated whenever risk preferences are preestablished as a matter of policy. For example, business entities may adopt expected values and risk neutrality as a norm for decision making. Also, the higher rate of estimated payments and payment-due filings among the five percent of tax filings with the highest adjusted gross income suggests a different pattern

for high income taxpayers than was observed among taxpayers in this study. However, the influence of a PR effect, even if confined to some middle portion of the 100,000,000 individual tax returns filed each year, still affects millions of taxpayers.

This picture implies that there is nothing to be gained in taxpayer satisfaction for the average taxpayer from changing the withholding system to reduce overpayments. It will be interesting to see whether or not taxpayers, despite withholding schedules designed to reduce overpayments, return to overpayment levels typical before 1986. Though there is little to be gained in taxpayer satisfaction, there may be losses in revenue from changes in the withholding system which reduce overpayments. If such changes increase the number of payment-due returns, taxpayers will report lower taxable incomes than if such changes had not occurred.

If this picture of the average taxpayer is accurate, the PR effect occurs within a general preference for risk avoidance. This means the taxpayer will pay to avoid risk. This result is commonplace in the application of the agency model to contracts where imposing risk on a risk-averse individual is costly, resulting in a second-best optimum. One of the results from this study is that in uncertain reporting situations, taxpayer subjects overpay taxes to avoid risk. In this experiment taxpayers preferred a high chance of overpaying taxes to increasing the chance of understating

their taxable income. A general implication for tax policy is that increasing the uncertainty about the correct reporting of an item will increase the amount reported, and decreasing the uncertainty will decrease the amount reported by risk-averse taxpayers. The extant empirical and analytical compliance literature is uniform in showing that increasing the risk or penalties of understating income increases the amount of taxable income reported by risk-averse taxpayers.

In general, reduction of uncertainty about tax outcomes has a potential for welfare improvements for risk-averse taxpayers by allowing them to avoid paying extra for risk avoidance. This is a general result which can be deduced from risk aversion alone, and it is in accord with the common wisdom that good tax systems have simple and certain outcomes. The PR effect simply means that any reductions in uncertainty have more value to a filer receiving a refund than the same filer who must make an additional payment. For example, if safe harbor rules reduce the uncertainty about the correct tax treatment of a transaction, then, all other things being equal, the safe harbor rules will be of more value to refund filers than to payment-due filers. To give yet another example, if the rapidity of change in tax laws adds to the uncertainty about the correct reporting of transactions, then slowing the rate of change in tax laws has more benefit for refund filers than for payment-due filers.

The implications of the PR effect for tax incentives parallels those for income and deductions. The more certain the taxpayer can be of the tax treatment of the transaction producing the benefit, the more value the risk-averse taxpayer is likely to obtain. A taxpayer filing for a refund will obtain more value from any reduction in uncertainty about the tax benefit than he will if he is filing with a payment due. Tax incentives also can affect the settlement amount. Since avoiding the payment due has more value than obtaining a refund of the same amount, any incentive allowing avoidance of the payment due has more incentive effect than if the taxpayer were in any case getting a refund. The decision to contribute to an Individual Retirement Account (IRA) can be deferred until the settlement amount is known. An implication of the PR effect is that taxpayers who have a payment due, and particularly taxpayers who can avoid a payment due by participating, have a stronger incentive to participate than taxpayers who were in a refund position without any IRA contribution (Feenberg and Skinner, 1989, p. 28).

The implication of the PR effect for enforcement is that the settlement amount should be a variable considered in enforcement strategies. The logic here is that payment-due filers prefer riskier filing choices and consequently have reported taxable incomes more likely to be too low than have refund filers. This is true for the average taxpayer who

prefers refunds, and it should hold as well for taxpayers who prefer to be underpaid at settlement.

The PR effect is triggered by uncertainty. The settlement amount is useful, therefore, only when uncertainty about the tax outcomes is important. In this experiment, the PR effect applied where there was uncertainty about the correct amount to report. Severe enforcement problems may exist where uncertainty is not a problem. For example, most taxpayers do not believe they should report in-kind income. Whether this comes from ignorance or belief in the legitimacy of ignoring that law, uncertainty does not seem to be a major factor. If uncertainty is not at issue, then, for example, searching for unreported in-kind income by selecting taxpayers with payments due at filing would be useless. But if there were some other variables available for selecting returns from taxpayers likely to have in-kind income, then of those returns, the payment-due filers can be expected to have taken riskier positions than they would have made if they were filing for refunds. Within any target group, the payment-due returns are likely to have riskier reporting choices than the refund returns.

The PR effect also has implications for tax practice. As with tax policy, there are general implications of the risk avoidance of average taxpayers, and the PR effect is an additional consideration. The removal of uncertainty about the proper reporting of items which the taxpayer is uncertain

about is a source of value to the taxpayer. In this experiment, subjects were willing to overpay taxes to avoid the risk of understating the income. Subjects also expressed agreement with the statement that taxpayers frequently do not know the correct amount to report for some items on their tax returns. The amount the taxpayer is willing to pay to avoid the risk of understating his income is a measure of the potential economic value to the taxpayer of certain knowledge of the legally correct treatment. Given highly risk-averse preferences and the difficulty of attaining confident knowledge of legally correct tax treatments, this amount would seem to be considerable. Reductions in uncertainty are a source of economic value to the average taxpayer which the tax preparer can deliver by knowing the correct amount of items to report.

By contrast, the potential value to a risk-averse taxpayer from taking a risky filing position where the correct tax treatment is subject to uncertainty is ambiguous. The dollar payoffs required to compensate the risk-averse taxpayer for assuming added risk are higher than expected from a risk neutral perspective. If tax practitioners are trained to value risky outcomes from a risk-neutral perspective, which is an emphasis compatible with formal academic business training, then they well may recommend riskier filing options than the average taxpayer prefers. They may also, from that perspective, fail to appreciate the

economic value to the risk-averse taxpayer of eliminating the uncertainty on items for which the tax expert knows there is little uncertainty. To the risk-averse taxpayer, one dollar saved by learning the proper treatment with near certainty is of much greater value than is one dollar saved by risking understatement.

The PR effect is an additional consideration. The risk preference of the taxpayer varies with the settlement. The value to the taxpayer of avoiding risk is lower if he has a payment due than if he is filing for a refund. Or, the average taxpayer will prefer riskier choices in a payment-due filing than in a refund filing. This is true both when the taxpayer has planned for a payment-due filing, in which case his selection of a framing is indicative of an interest in filing under a more risk-seeking frame, and when the taxpayer has planned to receive a refund. The average taxpayer seems to desire a frame of reference which is associated with risk-averse preferences, and he prefers highly risk-averse choices.

APPENDIX

EXPERIMENTAL INSTRUMENT

I. PURPOSE AND NATURE OF THE STUDY

The purpose of the project is to study personal judgments, so there are no right or wrong answers. Therefore, in each situation, choose the option you feel most comfortable with.

The session will last no longer than 60 minutes. All your answers will be anonymous.

Cash awards will be paid. The amount depends on your choices and on chance outcomes which will be resolved through the use of a game-type spinner. You may receive $10 or more, but you may receive nothing.

You are not required to participate in this study. You may withdraw if you wish.

II. CASH PAYOFFS

You will get a fund of money to spend to accomplish several tasks. The tasks are to pay income taxes, to buy a Health Care Plan for your family, and to select a gamble. Payments you make reduce your fund. Refunds and receipts increase your fund. Any of the fund not spent will be converted to cash awards at a rate of $.50 in cash for each $100 left over at the end of the session.

III. OVERVIEW OF THE TASKS

In the tax situations, the prepayments through withholding have already been made, and your task is to make any payment due, or to claim a refund, at the time of the tax filing. If the amount you choose to report turns out to under-report your tax, you must pay the taxes owed plus $300. The $300 is to covers all monetary and nonmonetary costs such as lost work time, professional service fees, anxiety over the deficiency, etc.

In the tax cases, the correct amount of tax is uncertain at the time you file your return. For example, your right to claim a dependency exemption for an adult child may be uncertain because of the facts of the case. To give another example, an activity which you consider a business may later be deemed to be a hobby, and hobby-related losses are not deductible. Or the uncertain item might be the gain on a home where the costs of the home sold are not documented. Please picture the uncertain item to be something for which the correct amount cannot be determined exactly at the time you file your return. You will choose how much (all, most, half, some, or none) of the uncertain item to report, and none of these choices would be illegal choices.

In the health care situation, the task is to purchase a family health care plan. You choose from among five family health care plans which differ in the health services covered. A family member may require services not covered by

the Plan you purchase. In that case you will have additional costs for the uncovered medical services needed by a member of your family.

In the gamble situation, you select one gamble from among five available choices. You may either win additional money, or you may have to make a payment from your fund.

IV. DETERMINING FUTURE OUTCOMES

In reality, the consequences of choices like these would only be known in the future. Here a spinner will be used to determine the unknown outcomes. The spinner exactly reflects the chances stated in the options. For example, in the tax situations Option B has a 15% risk of under-reporting your tax. If you choose Option B, you are determined to have under-reported only if the spinner lands in the first 15% of the wheel. Likewise, if you choose Option C, you are determined to have under-reported only if the spinner lands in the first 25% of the wheel. Option D has a 33% risk of under-reporting your liability, and Option E always has a 40% risk of under-reporting your liability.

The spinner is pictured here.

V. THE PROCEDURES

1. You receive a Package of Cases, a PAYMENT ENVELOPE, and a pack of money.
2. After reading a case, you circle the option you prefer.

(a) If that option requires a payment, you enclose the amount of the payment in the PAYMENT ENVELOPE. Then you go on to the next case.

(b) If your preferred option is a filing for a refund, you raise your hand to receive the amount of the refund. Then you go on to the next case.

3. After you have made your choices in all the cases, you will then spin the spinner for each case in which you chose an uncertain outcome.

(a) If the spin is favorable, you make no further payment.

(b) If the spin is unfavorable, you hand over the payment stated in the option.

4. After completing all the spins, any of your fund remaining will be converted to cash awards at the rate of $.50 for each $100 left over.

VI. SAMPLE CASE

A sample case appears on the next page.

SAMPLE CASE: DAY CARE INCOME

You are married filing jointly with taxable income of $25,780 not counting net income from providing day care in your home for several neighborhood pre-schoolers. The amount of this income you should include is uncertain because your costs for lunches, yard fencing, snacks, toys, use of your home, etc. are not documented.

You are now preparing your tax return. Your withholding totals $5,600. Consider both the risk of under-reporting and the dollar outcomes in choosing the amount of this income to report.

In Option A you include the total earnings. There is no risk of under-reporting your tax (though you may be overpaying). In Options C, C, D, and E, you include less and less in net income, and both your refund and also your chance of under-reporting increases. The risk of under-reporting your tax is 15% in Option B, 25% in Option C, 33% in Option D, and 40% in Option E.

TAX FILING OPTIONS -- Circle One

AMOUNT INCLUDED	(A) ALL	(B) MOST	(C) HALF	(D) SOME	(E) NONE
REFUND CLAIMED	REFUND $600	REFUND $700	REFUND $800	REFUND $900	REFUND $1,000
RISK & COST OF UNDERPAYMENT RISK TOTAL COST	0% $0	15% $200	25% $300	33% $400	40% $500

VII. TRIALS USING THE SAMPLE CASE

To include all of the Day Care Earnings, circle Option A, and raise you hand in order to be paid the $600 refund. There is no risk of under-reporting in this Option, and you would make no spin.

To include most of the Day Care Earnings, circle Option B and raise your hand to be paid the $700 refund. After making your choices in all the cases, you would then spin the spinner. If the spinner landed in the wedge with the letter B, you will be considered to have under-reported your true tax. You would have to pay $200 out of your fund. If the spinner landed anyplace else in the wheel, you would be considered to be <u>not</u> under-reported.

To include half of the Day Care Earnings, circle Option C and raise your hand to be paid the $800 refund. Later, you would spin the spinner. If the spinner landed in the wedge with the letter C, you will be considered to have under-reported your true tax. You would have to pay $300 out of your fund.

And so on for Options D and E.

JUDGMENT SITUATION PACKAGE

Identification Number_____.

In addition to this package, you should have a pack of money and a PAYMENT ENVELOPE.

Complete each case before you go on to the next one. Circle your choice and either raise your hand to receive a refund or enclose the payment in the PAYMENT ENVELOPE. Please do not look back at situations you have already marked unless you are asked to do so.

When you have completed this package, please raise your hand.

TAX CASE Y

The task is to file your income tax return. The correct amount of your income tax is uncertain because the correct amount of one deduction item is in doubt. You must choose the amount of this item to report. Your taxable income is $27,780 not counting this item. Your withholding is $5,400.

In Option A you deduct none of this item, and there is no risk of under-reporting your tax. In Options B, C, D, and E you deduct more and more of this item, and both your refund and also your chance of under-reporting increases. The cost if under-reported is unpaid taxes plus $300. The risk of under-reporting your tax is 15% in Option B, 25% in Option C, 33% in Option D, and 40% in Option E.

TAX FILING OPTIONS--Circle One

AMOUNT DEDUCTED	(A) NONE	(B) SOME	(C) HALF	(D) MOST	(E) ALL
REFUND CLAIMED	REFUND $400	REFUND $500	REFUND $600	REFUND $700	REFUND $800
RISK & COST OF UNDERPAYMENT					
RISK	0%	15%	25%	33%	40%
TOTAL COST	$0	$400	$500	$600	$700

Consider both the risk of under-reporting and the dollar amounts in choosing the amount of this item to deduct. Circle the Option you prefer, and raise your hand to receive your refund. Then go on to the next case.

TAX CASE Z

The task is to file your income tax return. The correct amount of your income tax is uncertain because the correct amount of one item of income is in doubt. You must choose the amount of this item to report. Your taxable income is $25,780 not counting this item. Your withholding is $4,800.

In Option A you include all this income, and there is no risk of under-reporting your tax. In Options B, C, D, and E, you include less and less of this income, and both your refund and your chance of under-reporting increases. The cost if under-reported is unpaid taxes plus $300. The risk of under-reporting your tax is 15% in Option B, 25% in Option C, 33% in Option D and 40% in Option E.

TAX FILING OPTIONS--Circle One

	(A)	(B)	(C)	(D)	(E)
AMOUNT INCLUDED	ALL	MOST	HALF	SOME	NONE
REFUND OR PAYMENT DUE	PAYMENT $200	PAYMENT $100	NONE $0	REFUND $100	REFUND $200
RISK & COST OF UNDERPAYMENT					
RISK	0%	15%	25%	33%	40%
TOTAL COST	$0	$400	$500	$600	$700

Consider both the risk of under-reporting and the dollar amounts in choosing the amount of this item to deduct. Circle the Option you prefer. If that Option requires a payment, enclose the payment in the PAYMENT ENVELOPE. If that Option has a refund, raise your hand to receive the refund. Then go on to the next case.

TAX CASE X

The task is to file your income tax return. The correct amount of your income tax is uncertain because the correct amount of one deduction item is in doubt. You must choose the amount of this item to report. Your taxable income is $27,780 not counting this item. Your withholding is $4,200.

In Option A you deduct none of this item, and there is no risk of under-reporting your tax. In Options B, C, D, and E you deduct more and more of this item, and your payment declines but your chance of under-reporting increases. The cost if under-reported is unpaid taxes plus $300. The risk of under-reporting your tax is 15% in Option B, 25% in Option C, 33% in Option D, and 40% in Option E.

TAX FILING OPTIONS--Circle One

AMOUNT DEDUCTED	(A) NONE	(B) SOME	(C) HALF	(D) MOST	(E) ALL
PAYMENT DUE	PAYMENT $800	PAYMENT $700	PAYMENT $600	PAYMENT $500	PAYMENT $400
RISK & COST OF UNDERPAYMENT					
RISK	0%	15%	25%	33%	40%
TOTAL COST	$0	$400	$500	$600	$700

Consider both the risk of under-reporting and the dollar amounts in choosing the amount of this item to deduct. Circle the Option you prefer, and enclose the payment due in the PAYMENT ENVELOPE. Then go on to the next case.

CHOOSE A GAMBLE

Your task is to choose one of the five gambles below. If you choose Option A, the outcome will be certain, and you will later pay $200 for sure. In Options B, C, D, and E you will receive either the more favorable outcome OR the less favorable outcome depending on the spinner results. In Options B, C, D, and E, the more favorable outcome is increasingly more favorable, but your chance of getting the more favorable payoff decreases.

GAMBLES -- Circle One

OPTIONS	CHANCE	WIN/ PAY	AMOUNT
(A)	100%	PAY	$200
(B)	85%	PAY	$100
	15%	PAY	$500
(C)	75%	PAY	$0
	25%	PAY	$500
(D)	67%	WIN	$100
	33%	PAY	$500
(E)	60%	WIN	$200
	40%	PAY	$500

Circle your preferred gamble. Then go on to the next case.

HEALTH CARE PLAN

The task is to purchase a Health Care Plan for your family. Your employer pays for part of the plan, and you pay the rest. You choose from among five plans which differ in the medical services covered. Plan A is the full coverage policy. In Options B, C, D and E, you buy Plans which have fewer and fewer covered services. If no one in your family needs any of the uncovered services, then the cheaper Plans provide coverage as good as that of Plan A.

If you choose Plan A, there is no risk of additional costs, but Plan A costs you more. In Options B, C, D, and E your net cost for the Health Plan decreases, but the chance that someone in your family will need uncovered medical services (costing you additional money) increases. The risk of added cost is 15% in Option B, 25% in Option C, 33% in Option D, and 40% in Option E.

HEALTH CARE OPTIONS--Circle One

HEALTH PLANS	(A) PLAN A	(B) PLAN B	(C) PLAN C	(D) PLAN D	(E) PLAN E
YOUR COST	$800	$700	$600	$500	$400
RISK AND COST OF UNCOVERED SERVICES					
RISK	0%	15%	25%	33%	40%
YOUR COST	$0	$400	$500	$600	$700

The chart shows your cost as well as the risk and cost of needing uncovered services for each Health Care Plan. Circle the Plan you prefer, and put the amount of your cost in the PAYMENT ENVELOPE. Then go on to the next case.

ADDITIONAL RATING OF JUDGMENT SITUATIONS

On this page, please prepare an additional rating for these five Judgment Situations. For this rating please use each Option (A, B, C, D, and E) once and only once. Assign the A (the least risky option) to the Judgment Situation for which you are least willing to take risk, and so on. This means you should Assign the E (the most risky option) to the Judgment Situation for which you are willing to take the most risk.

Please look back at all the Judgment Situations and compare them to make this additional rating.

USE EACH OPTION (A,B,C,D AND E) ONCE AND ONLY ONCE

SITUATION	OPTION
TAX CASE X	_____
CHOOSE A GAMBLE	_____
TAX CASE Y	_____
BUY HEALTH PLAN	_____
TAX CASE Z	_____

[WHEN YOU HAVE COMPLETED THIS ADDITIONAL RATING, PLEASE RAISE YOUR HAND]

REFERENCES

Allingham, M.G., and A. Sandmo, "Income Tax Evasion: A Theoretical Analysis," The Journal of Public Economics, 1 (1972), 323-338.

Arrow, K.J., Essays in the Theory of Risk Bearing (Chicago: Markham Pub. Co., 1971).

Baiman, S., and B.L. Lewis, "An Experiment Testing the Behavioral Equivalence of Strategically Equivalent Employment Contracts," Working Paper (Carnegie Mellon University, November, 1987).

Becker, G.S., "Crime and Punishment: An Economic Approach," Journal of Political Economy, 76 (March-April, 1968), 169-217.

Berg, J.E., and J. Dickhaut, "Doctoral Consortium: Questions and Preliminary Answers on Experimental Economics," Working Paper (University of Minnesota, August, 1986).

Berg, J.E., L.A. Daley, J.W. Dickhaut and J. O'Brien, "Moral Hazard and Risk Sharing: Experimental Evidence," Working Paper (University of Minnesota, August, 1987).

Bierman, H. Jr., and S. Smidt, The Capital Budgeting Decision, 3rd ed. (New York: Macmillan, 1971).

Chang, O.H., D.R. Nichols, and J.J. Schultz, "Taxpayer Attitudes Toward Tax Audit Risk," The Journal of Economic Psychology, 8 (1987), 299-309.

Conner, E., "Crime Commissions and Criminal Procedure in the United States Since 1920: A Bibliography, January, 1920 to June, 1927," Journal of Criminal Law and Criminology, 21 (1930), 129-144.

Davis, J.S., and C. Swenson, "The Role of Experimental Economics in Tax Policy Research, The Journal of

the American Taxation Association, 9 (Fall, 1988), 40-59.

Dean, J., Capital Budgeting (New York: Columbia University Press, 1951).

Doernberg, R.L., and F. S. McChesney, "On the Accelerating Rate and Decreasing Durability of Tax Reform," Minnesota Law Review 71 (1987), 913-962.

Duesenberry, J.S., Income, Saving and the Theory of Consumer Behavior (Cambridge: Harvard University Press, 1949).

Duncan, W.A., D. LaRue and P.M.J. Reckers, "An Empirical Examination of the Influence of Selected Economic and Noneconomic Variables on Decision Making By Tax Professionals," in S.M. Jones ed., Advances in Taxation (Greenwich, Connecticut: JAI Press, 1989), 1 (1989), 91-106.

Einhorn, H.J., and R.M. Hogarth, "Decision Making Under Ambiguity," Journal of Business, 59 (1986), S225-S250.

Elffers, H., R.H. Weigel and D.J. Hessing, "The Consequences of Different Strategies For Measuring Tax Evasion Behavior," The Journal of Economic Psychology, 8 (1987), 311-337.

Ellwood, C.A., "Scientific Methods of Studying Human Society," Journal of Social Forces, 2 (1924), 328-332.

Feenberg, D. and J. Skinner, "Sources of IRA Saving," in L.H. Summers ed., Tax Policy and the Economy 3, National Bureau of Economic Research, (Cambridge, MIT Press, 1989), 25-46.

Fisher, I., The Theory of Interest (New York: Macmillan, 1930).

Friedman, M. A Theory of the Consumption Function (Princeton, N.J.: Princeton University Press, 1957).

Friedman, M., and L.J. Savage, "The Utility Analysis of Choices Involving Risks," Journal of Political Economy, 56 (1948), 279-304.

Gravelle, J.G., and L.B. Lindsey, "Capital Gains," Tax Notes, 38 (January 25, 1988), 397-405.

Harwood, G., J. Pate and A. Schneider, "Framing Bias in Managerial Accounting Decisions," Working Paper, (Georgia State University, March, 1988).

Hessing, D.J., H. Elffers and R.H. Weigel, "Exploring the Limits of Self Reports and Reasoned Action: An Investigation of the Psychology of Tax Evasion Behavior," The Journal of Personality and Social Psychology, 54 (1988), 405-413.

Hite, P.A., "The Effect of Peer Reporting Behavior on Taxpayer Compliance," Journal of the American Taxation Association, 9 (Spring, 1988), 47-64.

Hite, P.A., B.R. Jackson and M.W. Spicer, "The Effects of Decision Framing on Taxpayer Compliance," Working Paper (July, 1988).

Holmstrom, B., "Moral Hazard and Observability," The Bell Journal of Economics, 10 (Spring, 1979), 74-91.

Homuth, H.G., "The Taxpayer Angle on Local Income Tax Administration," Chapter 39 of Tax Institute, Income Tax Administration (New York: Tax Institute, 1948), 329-349.

Internal Revenue Service, Statistics of Income (Washington, D.C.: U.S. Government Printing Office, various years).

Jackson, B.R., and S.M. Jones, "Salience of Tax Evasion Penalties Versus Detection Risk," The Journal of the American Taxation Association, 6 (Spring, 1985), 7-17.

Jackson, B.R., and V.C. Milliron, "Tax Compliance Research: Findings, Problems and Prospects," The Journal of Accounting Literature, 5 (1986), 125-165.

Johnson, M.F., "Service Reissues Simplified Form W-4," Tax Notes, 34 (March 9, 1987), 946-947.

Jones, E.P., "Press Watch," Tax Notes, 33 (December 8, 1986), 971-972.

Kahn, B.E., and R.K. Sarin, "Modeling Ambiguity in Decisions Under Uncertainty," Journal of Consumer Research, 15 (1988), 265-271.

Kahneman, D., "Rational Choice and the Framing of Decisions," Journal of Business, 59 (October, 1986), S251-S278.

Kahneman, D., J.L. Knetsch and R. Thaler, "Fairness as a Constraint on Profit Seeking: Entitlements in the Market," American Economic Review, 76 (1986), 728-741.

Kahneman, D., and A. Tversky, "Prospect Theory: An Analysis of Decision Under Risk," Econometrica, 47 (March, 1979), 263-291.

Kahneman, D., and A. Tversky, "The Psychology of Preferences," Scientific American, 249 (January, 1982), 160-173.

Kay, J.A., and M.A. King, The British Tax System, 4th ed. (Oxford: Oxford University Press, 1986).

Keppel, G., Design and Analysis (Englewood Cliffs, New Jersey: Prentice-Hall, 1982).

Keren, G.B., and J. G. W. Raaijmakers, "On Between-Subjects versus Within-Subjects Comparisons in Testing Utility Theory," Organization Behavior and Human Decision Processes, 41 (1988), 233-247.

Kotlarski, M., "Capital Gains and Tax Policy," Tax Notes, 41 (October 17, 1988), 319-327.

Kuhlman, A.F., A Guide to Material On Crime and Criminal Justice (New York: H.W. Wilson Co., 1929).

LaTour, S.A., and P.W. Miniard, "Considerations in the Analysis of Repeated Measures Designs," Working Paper (University of Florida, December, 1978).

Lemert, E.M., Human Deviance, Social Problems, and Social Control (Englewood Cliffs, N.J.: Prentice-Hall, 1964).

Levin, I.P., R.D. Johnson and M.L. Davis, "How Information Frame Influences Risky Decisions," The Journal of Economic Psychology, 8 (1987), 43-54.

Lindholm, R.W., "General Appraisal of State Income Tax Administrative Problems," Chapter 33 of Tax Institute, Income Tax Administration (New York: Tax Institute, 1948), 263-272.

Long, S.B., and J.A. Swingen, "An Approach to the Measurement of Tax Law Complexity," The Journal of the American Taxation Association, 8 (Spring, 1987), 22-36.

Luce, R.D., and H. Raiffa, Games and Decisions (New York: John Wiley, 1967).

Lutz, F., and V. Lutz, The Theory of the Investment of the Firm (Princeton, N.J.: Princeton University Press, 1951).

Machina, M.J., "The Economic Theory of Individual Behavior Toward Risk: Theory, Evidence and New Directions," Technical Report No. 433 (Stanford University Institute for Mathematical Studies in the Social Sciences, October, 1983).

Machina, M.J., "Choice Under Uncertainty: Problems Solved and Problems Unsolved," Economic Perspectives, 1 (1987), 121-154.

Madeo, S.A., A. Schepanski, and W.C. Uecker, "Modeling Judgments of Taxpayer Compliance," The Accounting Review, 62 (1987), 323-342.

Mead, J., "The Lock-In Effect of Capital Gains Taxation and the Willingness to Undertake Risky New Investment, "Working Paper (University of Houston, June, 1988).

Merton, R.K., "Social Structure and Anomie," American Sociological Review, 3 (October, 1938), 672-682.

Milliron, V.C., "An Analysis of the Relationship Between Tax Equity and Tax Complexity," The Journal of the American Taxation Association, 6 (Fall, 1985a), 19-33.

Milliron, V.C., "A Behavioral Study of the Meaning and Influence of Tax Complexity," The Journal of Accounting Research, 23 (1985b), 794-816.

Mirrlees, J.A., "An Exploration in the Theory of Optimum Income Taxation," Review of Economic Studies, 38 (1971), 175-208.

Mirrlees, J.A., "The Optimal Structure of Incentives and Authority Within An Organization," Bell Journal of Economics, 7 (1976), 105-131.

Moore, M., B. Steece and C. Swenson, "Some Empirical Evidence On Taxpayer Rationality," The Accounting Review, 60 (1985), 18-32.

National Commission on Law Observance and Enforcement, Reports, 6 vols. (Washington, D.C.: U.S. Government Printing Office, 1931).

Odum, H.W. and K. Jocher, An Introduction to Social Research (New York: H. Holt, 1929).

Payne, J.W., D.J. Laughhunn, and R. Crum, "Multiattribute Risk Choice Behavior: The Editing of Complex Prospects," Management Science, 30 (1984), 1350-1363.

Plott, C.R., "Industrial Organization Theory and Experimental Economics," Review of Economic Literature, 20 (1982), 1485-1527.

Plott, C., "Rational Choice in Experimental Markets," Journal of Business, 59 (1986), S301-S327.

Polanyi, M., Scientific Thought and Social Reality (New York: International Universities Press, 1974).

Pratt, J.W., "Risk Aversion in the Small and in the Large," Econometrica, 32 (1964), 122-136.

Quinney, R., Criminology (Boston: Little Brown, 1975).

Reckless, W.C., Six Boys in Trouble, A Sociological Case Book (Ann Arbor: Edwards Brothers, 1929).

Reinganum, J., and L.L. Wilde, "Income Tax Compliance in a Principal Agent Framework," The Journal of Public Economics, 26 (1985), 1-18.

Reinganum, J., and L.L. Wilde, "Tax Practitioners and Tax Compliance," Social Science Working Paper 666 (California Institute of Technology, March, 1988).

Robben, H.S.J., P. Webley, H. Elffers and D.J. Hessing, "Investigating Tax Evasion in Prospect Theoretical Terms: Evidence From Experimental and Questionnaire Data," Annual Meeting of Law and Society, Vail, Colorado (June, 1988).

Rothschild, M., and J.E. Stiglitz, "Increasing Risk I: A Definition," Journal of Economic Theory, 2 (1970), 225-243.

Schadewald, M., "Reference Point Effects in Taxpayer Decision Making," The Journal of the American Taxation Association, 10 (Spring, 1989), 68-84.

Schepanski, A., and Kelsey, D., "Testing for Framing Effects in Taxpayer Compliance," Working Paper (University of Iowa, 1988).

Sellin, T., Culture Conflict and Crime (New York: Social Science Research Council, 1938).

Shaw, C., The Jack-Roller (Chicago: University of Chicago Press, 1930).

Shaw, C., F.M. Zorbaugh, H.D. McKay and L.S. Cottrell, Delinquency Areas (Chicago: University of Chicago Press, 1929).

Shefrin, H., and M. Statman, "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence," The Journal of Finance, 40 (1985),777-792.

Sheppard, L.A., "Chief Counsel Nelson Comments on Form W-4," Tax Notes, 34 (February 2, 1987), 439-440.

Smith, K. W. and K. A. Kinsey, "Understanding Taxpayer Behavior:A Conceptual Framework With Implications for Research," Law and Society, 21 (1987), 501-524.

Spicer, M.W., "Civilization at a Discount: The Problem of Tax Evasion," The National Tax Journal, 39 (1986), 13-20.

Spicer, M.W. and L.A. Becker, "Fiscal Inequity and Tax Evasion: An Experimental Approach," The National Tax Journal, 33 (1980), 171-175.

Sutherland, E.H., and D.R. Cressey, Criminology, 8th ed. (Philadelphia: J.B. Lippincott, 1970).

Teuber, J., "Gibbs Says IRS Will Decide Whether To Replace Form W-4 By March 1," Tax Notes, 34 (February 16, 1987), 633-635.

Thaler, R., "Toward A Positive Theory of Consumer Choice," Journal of Economic Behavior and Organization, 1 (March, 1980), 39-60.

Thaler, R., "Mental Accounting and Consumer Choice," Marketing Science, 4 (Summer, 1985), 199-214.

Thomas, W.I., The Unadjusted Girl (Boston: Little Brown, 1923).

Thomas, W.I., and Znaniecki, F., The Polish Peasant in Europe and America, 2 vols. (New York: A.A. Knopf, 1927).

Thrasher, F.M., The Gang (Chicago: University of Chicago Press, 1927).

Toder, E. and L. Ozanne, "CBO Works on Capital Gains," Tax Notes, 39 (June 20, 1988), 1441-1453.

Tversky, A., "Elimination By Aspects: A Theory of Choice," Psychological Review, 79 (1972), 281-299.

Tversky, A., and D. Kahneman, "The Framing of Decisions and the Psychology of Choice," Science, 211 (January, 1981), 453-458.

Uhlfeder, M.N., "IRS Unveils New Form W-4," Tax Notes, 33 (November 25, 1986), 703.

Von Neumann, J., and O. Morgenstern, Theory of Games and Economic Behavior (Princeton: Princeton University Press, 1944).

Weigel, R.H., D.J. Hessing and H. Elffers, "Tax Evasion Research: A Critical Appraisal and Theoretical Model," The Journal of Economic Psychology, 8 (1987), 215-235.

Westat Inc., Individual Income Tax Factors Study: Qualitative Research Results, Prepared for the Internal Revenue Service (Rockville, Maryland: author, February, 1980).

Witte, A.D., and D.F. Woodbury, "The Effect of Tax Laws and Tax Administration on Tax Compliance: The Case of the U.S. Individual Income Tax," The National Tax Journal, 38 (1985), 1-13.

Witte, J.F., The Politics and Development of the Federal Income Tax (Madison, Wisconsin: University of Wisconsin Press, 1985).

Zorbaugh, H.W., Gold Coast and the Slum (Chicago: University of Chicago Press, 1929).

BIOGRAPHICAL SKETCH

Richard Dusenbury was born in the upper Midwest and grew up in Gainesville, Florida. He attended public schools and the University of Florida where he earned a bachelor's and master's degree in history. He also obtained a Ph.D. in history from the University of Wisconsin-Madison. As historian, he taught at Simpson College and wrote on the history of behavioral sciences, myth in United States history, and United States forest history.

Dusenbury began a career in accounting and finance with a position of Trainer attached to the Finance Division of Blue Cross and Blue Shield of Florida, where he was promoted to increasingly responsible positions managing various financial analysis and reporting functions within a controllership career track. He earned M.B.A. and Master of Accountancy degrees and became a C.P.A. during these years.

As the demands of his career increased, Dusenbury realized that he was not motivated by either wealth or power, the two chief rewards of executive status. Having been a single parent for some years, he wanted to spend more time

with his daughters. He therefore took a position at Flagler College where he implemented and staffed an undergraduate accounting curriculum and briefly chaired the Department of Economics and Business. In 1989, he accepted an academic appointment in accounting at Florida State University.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Sandra S. Kramer, Chairperson
Associate Professor,
Fisher School of Accounting

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Jon H. Hamilton
Assistant Professor of Economics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

John L. Kramer
Arthur Andersen Professor,
Fisher School of Accounting

This dissertation was submitted to the Graduate Faculty of the Fisher School of Accounting in the College of Business Administration and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree Doctor of Philosophy.

December, 1989 Dean of the Graduate School